# Journal of ATMOSPHERIC and TERRESTRIAL PHYSICS

EDITORIAL ADVISORY BOARD

H. O. G. ALFVÉN–STOCKHOLM · J. BARTELS–GÖTTINGEN · W. J. G. BEYNON–ABERYSTWYTH
P. M. S. BLACKETT–LONDON · H. G. BOOKER–ITHACA, N.Y. · E. G. BOWEN–SYDNEY
SIR DAVID BRUNT–LONDON · R. BUREAU–BAGNEUX · S. CHAPMAN–ALASKA
J. EGEDAL–COPENHAGEN · V. F. HESS–NEW YORK · E. O. HULBURT–WASHINGTON, D.C.
H. ISRAEL–AACHEN · D. K. MENZEL–CAMBRIDGE, MASS. · S. K. MITRA–CALCUTTA
M. NICOLET–UCCLE · E. H. NORINDER–UPPSALA · J. PATON–EDINBURGH
N. V. PUSHKOV–MOSCOW · J. A. RATCLIFFE–SLOUGH · O. E. H. RYDBECK–GOTHENBURG
J. C. W. SCOTT–OTTAWA · SIR GEORGE SIMPSON–LONDON · H. SOLBERG–OSLO
J. VELDKAMP–DE BILT · M. WALDMEIER–ZÜRICH

EDITOR-IN-CHIEF

SIR EDWARD APPLETON, F.R.S.
EDINBURGH

VOL.
20
1961

VOLUME 20
1961

PERGAMON PRESS
OXFORD · LONDON · NEW YORK · PARIS

VOL.
20
1961

# CONTENTS

## No. 1.


W. G. MATHEWS and L. WALLACE: Auroral temperature determination from $A^2\Pi \rightarrow X^2\Sigma$ bands of the $N_2^+$ molecule . . . . . 1

S. A. BOWHILL: The scattering of radio waves by an extended randomly refracting medium 9

S. A. BOWHILL: The effective recombination coefficient of an ionosphere containing a mixture of ions . . . . . 19

J. HEADING: Analytical considerations of ionospheric windows for low frequency radio waves . . . . . 31

J. WHITE, N. A. PORTER and C. D. LONG: Observations on Čerenkov radiation from the night sky . . . . . 40

J. D. WHITEHEAD: The formation of the sporadic-*E* layer in the temperate zones . . 49

S. GRUBER: Statistical analysis of radio star scintillation . . . . . 59

Research Notes

A. H. SHAPLEY: International Geophysical Calender for 1961 . . . . . 72

T. M. DONAHUE: A note on polar black outs . . . . . 76

R. MÜHLEISEN: Electrode effect measurements above the sea . . . . . 79

B. N. BHARGAVA and R. V. SUBRAHMANYAN: Geomagnetic disturbance effects in equatorial $E_s$ . . . . . 81


## Nos. 2/3


L. WALLACE: Seasonal variation and interpretation of the OH rotational temperature of the airglow . . . . . 85

C. M. MINNIS: An estimate of the peak sunspot number in 1968 . . . . . 94

ROSEMARY HUTTON and R. W. H. WRIGHT: Diurnal variation of earth currents at the equator . . . . . 100

S. N. GHOSH and A. SHARMA: Excitation of green and red lines in the night airglow . 110

R. L. DOWDEN: A theoretical model of electron density distribution along a geomagnetic line of force in the exosphere . . . . . 122

C. D. WATKINS: Auroral radio-echoes and magnetic disturbances . . . . . 131

C. D. WATKINS: Temporal variations of auroral radio-echo activity in sub-auroral latitudes 140

A. J. REYNOLDS: Energy transport due to turbulent mixing in the atmosphere . . 149

J. K. HARGREAVES: Random fluctuations in very low frequency signals reflected obliquely from the ionosphere . . . . . 155

J. P. DOUGHERTY: On the influence of horizontal motion of the neutral air on the diffusion equation of the *F*-region . . . . . 167

M. HERTZBERG: Ion–neutral reactions . . . . . 177

J. E. C. GLIDDON and P. C. KENDALL: Theoretical world curves of maximum *F2* ionization—II . . . . . 183

H. UNZ: On the evaluation of the group refractive index in case of no collisions . . 189

M. C. PANDE and S. S. VARMA: Study of variations of the forbidden oxygen lines in the nightglow . . . . . 195

R. K. MACCRONE and F. R. N. NABARRO: Behaviour of broadcast frequency waves at oblique incidence during an annular eclipse . . . . . 200

Research Notes

H. G. MARTIN: Wave-guide mode propagation of very low frequency radio waves . . 206

J. TITHERIDGE: The analysis of night-time $h'(f)$ records . . . . . 209

M. YAMAMOTO and H. MAEDA: The simultaneity of geomagnetic sudden impulses . . 212

J. FRIHAGEN and J. TRÖIM: On the large scale regions of irregularities producing scintillation of signals transmitted from earth satellites . . . . . 215

A. OMHOLT: On the excitation of the First Positive $N_2$ bands in aurorae . . . . 217

Book Reviews

"Physics and Medicine of the Atmosphere and Space". Edited by O. O. BENSON, Jr. and H. STRUGHOLD. (Reviewed by Sir EDWARD APPLETON) . . . . . 220

"Atlantic Hurricanes" by G. E. DUNN and B. I. MILLER. (Reviewed by J. PATON) . 220

"Variational Principles in Dynamics and Quantum Theory" by W. YOURGRAU and S. MANDELSTAM. (Reviewed by R. M. SILLITTO) . . . . . 221


VOL. 20 1961

"Turbulent Transfer in the Lower Atmosphere" by C. H. B. PRIESTLEY. (Reviewed by D. H. MCINTOSH) . . . . . 221
"Space Research—Proceedings of the First International Space Symposium". Edited by H. KALLMAN BIJL. (Reviewed by Sir EDWARD APPLETON) . . . . . 222
"Electromagnetic Wave Propagation". Edited by M. DESIRANT and J. L. MICHIELS. (Reviewed by Sir EDWARD APPLETON) . . . . . 222
*K*-indices of geomagnetic activity for September–December 1960 . . . . . 224
Papers to be published in future issues . . . . . 228

**No. 4**

G. F. FOOKS and I. L. JONES: Correlation analysis of the fading of radio waves reflected vertically from the ionosphere . . . . . 229
J. HARWOOD: Some observations of the occurrence and movement of sporadic-*E* ionization 243
H. RISHBETH and C. S. G. K. SETTY: The *F*-layer at sunrise . . . . . 263
H. RISHBETH: Diffusion of ionization in the sunrise *F*-layer . . . . . 277
K. S. RAJA RAO: On the seasonal variation in lunar and solar geomagnetic tides in the geomagnetic equatorial region . . . . . 289
Research Notes
F. H. NORTHOVER: The effect of multiple atmospheric inversions on tropospheric radio propagation . . . . . 295
E. BHAGIRATHA RAO and B. RAMACHANDRA RAO: Travelling disturbances in the *F*-region over Waltair . . . . . 296
W. R. PIGGOTT and L. W. BARCLAY: The reflection of radio waves from an iceshelf . 298
Letter to the Editor
G. C. REID: Discussion of the paper "Observed variations in the amplitude scintillations of the Cassiopeia (23N5A) source" by H. J. A. Chivers . . . . . 300

VOL. 20 1961

## AUTHOR INDEX

(B.R.) denotes a book review

APPLETON E. V. (B.R.'s) 220, 222
BARCLAY L. W. 298
BHARGAVA B. N. 81
BOWHILL S. A. 9, 19
DONAHUE T. M. 76
DOUGHERTY J. P. 167
DOWDEN R. L. 122
FOOKS G. F. 229
FRIHAGEN J. 215
GHOSH S. N. 110
GLIDDON J. E. C. 183
GRUBER S. 59
HARGREAVES J. K. 155
HARWOOD J. 243
HEADING J. 31
HERTZBERG M. 177
HUTTON R. 100
JONES I. L. 229
KENDALL P. C. 183
LONG C. D. 40
MACCRONE R. K. 200
MAEDA H. 212
MARTIN H. G. 206
MATHEWS W. G. 1
MCINTOSH D. H. (B.R.) 221
MINNIS C. M. 94
MÜHLEISEN R. 79
NABARRO F. R. N. 200
NORTHOVER F. H. 295
OMHOLT A. 217
PANDE M. C. 195
PATON J. (B.R.) 220
PIGGOTT W. R. 298
PORTER N. A. 40
RAO B. R. 296
RAO E. B. 296
RAO K. S. R. 289
REID G. C. 300
REYNOLDS A. J. 149
RISHBETH H. 263, 277
SETTY C.S.G.K. 263
SHAPLEY A. H. 72
SHARMA A. 110
SILLITTO R. M. (B.R.) 221
SUBRAHMANYAN R. V. 81
TITHERIDGE J. 209
TRÖIM J. 215
UNZ H. 189
VARMA S. S. 195
WALLACE L. 1, 85
WATKINS C. D. 131, 140
WHITE J. 40
WHITEHEAD J. D. 49
WRIGHT R. W. H. 100
YAMAMOTO M. 212

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 1 to 8. Pergamon Press Ltd. Printed in Northern Ireland

# Auroral temperature determination from $A\ ^2\Pi \rightarrow X\ ^2\Sigma$ bands of the $N_2^+$ molecule*

W. G. Mathews and L. Wallace
Yerkes Observatory, University of Chicago, Williams Bay, Wisconsin

(*Received* 16 *July* 1960)

**Abstract**—A comparison of observed $N_2^+ A\ ^2\Pi \rightarrow X\ ^2\Sigma(2,0)$ and (3,1) band profiles with calculated model intensity profiles at various temperatures indicates that the auroral rotational temperature is $350 \pm 75°$K. This temperature is higher than the average of the determinations from the $N_2^+ B\ ^2\Sigma \rightarrow X\ ^2\Sigma$ bands but the difference is not sufficiently significant to affect the hypothesis that both band systems have a common excitation mechanism.

## 1. Introduction

The $N_2^+ A\ ^2\Pi \rightarrow X\ ^2\Sigma$ bands in the near infra-red were first identified in an auroral spectrum by Meinel (1951). Shortly afterwards, Dalby and Douglas (1951) succeeded in producing these bands in the laboratory and later Douglas (1953) analysed their fine structure and determined rotational constants for several of the bands.

The twofold purpose of the present work is to determine the sensitivity of the intensity profiles of the $A \rightarrow X$ bands to rotational temperature and, secondly, with the aid of auroral observations of the $A \rightarrow X$ bands, to determine $A \rightarrow X$ rotational temperatures for comparison with the $B \rightarrow X$ (or first-negative) temperatures recently found by a number of investigators. The work of Roesler *et al.* (1958) indicates that the $B \rightarrow X$ temperatures are probably kinetic temperatures, but this interpretation is still inconclusive and it may be the case that the observed rotational temperature depends to some extent on the excitation mechanism. If the excitation mechanisms are the same, the $A \rightarrow X$ and $B \rightarrow X$ temperatures should agree.

As will be seen, in determining $T_{rot}$ for the $A \rightarrow X$ bands the doublet nature of the $^2\Pi$ level requires a somewhat more complicated analysis than the $^2\Sigma \rightarrow {}^2\Sigma$ case in which there are, for the purpose of rotational temperature determination, only two branches compared with eight effective branches in the transition considered here (see Fig. 1). Furthermore, $^2\Sigma$ levels are always strictly Hunds coupling case (b) whereas $^2\Pi$ levels pass from case (a) to case (b) as $K'$ increases, making the determination of the rotational part of the transition probability more complicated. In addition to these complications, the rotational structure of the $^2\Pi \rightarrow {}^2\Sigma$ bands has not been resolved on the auroral spectrograms.

## 2. Observational Material

Many auroral spectrograms taken with the 9 in. aperture Meinel spectrograph show the $^2\Pi \rightarrow {}^2\Sigma$ bands, but only three show them in strong contrast to the air-

* The research reported in this paper was supported in part by the Geophysics Research Directorate of the Air Force Cambridge Research Center under Contract AF 19(604)-3044.

glow background. These observations, having a dispersion of about 70 Å/mm, were obtained on 10 January, 1/2 August and 3 October, 1959. Of these, the latter two show only the (2,0) band and the first shows also the (3,1) band, all other bands being of lower intensity and strongly contaminated with night-sky emission. These four bands were converted to an intensity scale with the use of characteristic curves constructed from additional spectrograms exposed through an optical step wedge. The observed profiles are given in Fig. 2. The positions of the principal OH contaminations are also indicated.

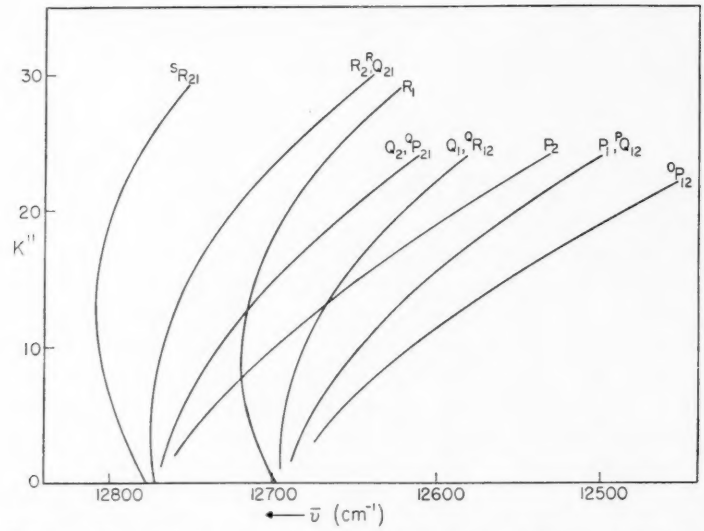


Fig. 1. A Fortrat diagram of (2,0) band based on Douglas' (1953) analysis.

## 3. Model Calculations

With the rotational constants given by Coster and Brons (1931), the term values for the $^2\Sigma$ state are given by

$$F(K'') = 1{\cdot}920K''(K'' + 1) - 2{\cdot}9 \times 10^{-6}K''^2(K'' + 1)^2.$$

The $^2\Pi$ term values for any amount of coupling intermediate between Hund's cases (a) and (b) are, according to Herzberg (1950),

$$F_1(J') = B_{v'}\{(J' + \tfrac{1}{2})^2 - 1 - \tfrac{1}{2}[4(J' + \tfrac{1}{2})^2 + \lambda_{v'}(\lambda_{v'} - 4)]^{\frac{1}{2}}\} - D_{v'}J'^4$$

and

$$F_2(J') = B_{v'}\{(J' + \tfrac{1}{2})^2 - 1 + \tfrac{1}{2}[4(J' + \tfrac{1}{2})^2 + \lambda_{v'}(\lambda_{v'} - 4)]^{\frac{1}{2}}\} - D_{v'}(J' + 1)^4$$

where, for the (2,0) band $B'_2 = 1{\cdot}695$, and $D'_2 = 4 \times 10^{-6}$, according to Douglas (1953).

It should be pointed out that Douglas' values of the coupling constants $A_v$, determined from the parameter $\lambda_{v'}(\lambda_{v'} - 4)$, are in error. In the solutions of this

expression for $\lambda_{v'}$, Douglas took the positive roots instead of the negative. The correct values of $A_v$ obtained from the negative roots are

$$A_2 = -74{\cdot}62$$
$$A_3 = -74{\cdot}60$$
$$A_4 = -74{\cdot}58$$

which are in much better agreement with the value $A = -75$, suggested by MEINEL (1951). It follows then that $\lambda_2 = -44{\cdot}02$. [We have used MEINEL's (1951) revised vibrational numbering which was used by Douglas and subsequently verified by LIU (1959).]

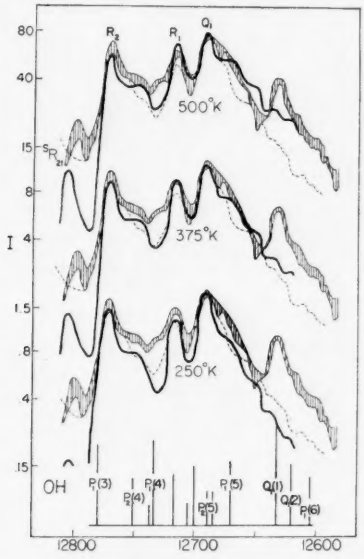


Fig. 2. In each of the three diagrams the model profile, shown as a solid heavy line, is compared to the observations represented by solid lines and a hatched region which is the same in each of the three cases. The hatched area is the envelope of the three superimposed observations of the (2,0) band, and the dotted lines represent the single (3,1) observation only where it differs significantly from the shaded region. The ordinates are logarithmic and the abscissae refer only to the (2,0) band. In each case the ordinate was varied for the best possible fit. The OH spectrum is represented schematically at the bottom of the diagram. The $P$ lines are part of the (9,4) band; the remainder are due to the (5,1) band where the unlabelled features are part of the $R$ branches.

With a knowledge of these energy levels it was possible to complete and extend Douglas' list of line wavenumbers to values of $K''$ required by the temperatures selected for the trial intensity models. This extended list is given in Table 1.

Because this transition is permitted and the lifetime in the $^2\Pi$ state is short compared with the interval between collisions it would seem extremely unlikely that thermal equilibrium could be established in the $^2\Pi$ state. Furthermore, it is easy to eliminate this possibility by using the observed profiles.

If it is assumed momentarily that equilibrium is attained in the upper ($^2\Pi$) state, the lower-lying *F1*-levels can be expected to be, in general, more highly

Table 1. Wavenumbers followed by intensity factors for the (2,0) band

| $K''$ | $(P_1 + {}^{P}Q_{12})$ | $(Q_1 + {}^{Q}R_{12})$ | $R_1$ | $P_2$ | $(Q_2 + {}^{Q}P_{21})$ | $(R_2 + {}^{R}Q_{21})$ | ${}^{S}R_{21}$ | ${}^{O}P_{12}$ |
|---|---|---|---|---|---|---|---|---|
| 0 | | | 12,698·7 0·521 | | | 12,771·8 0·333 | 12,777·0 0·146 | |
| 1 | | 12,694·9 0·913 | 703·2 0·651 | | 12,768·2 0·500 | 73·1 0·687 | 81·8 0·249 | |
| 2 | 12,687·2 0·479 | 95·5 1·322 | 07·1 0·804 | 12,760·3 0·167 | 65·5 0·855 | 74·1 1·035 | 86·2 0·339 | |
| 3 | 84·0 0·849 | 95·5 1·74 | 10·6 0·972 | 53·9 0·313 | 62·6 1·251 | 74·7 1·374 | 90·4 0·417 | 12,675·7 0·087 |
| 4 | 80·2 1·196 | 95·2 2·16 | 13·5 1·151 | 47·3 0·465 | 59·4 1·66 | 75·0 1·70 | 94·0 0·485 | 68·7 0·178 |
| 5 | 75·9 1·528 | 94·2 2·59 | 15·8 1·340 | 40·2 0·625 | 55·8 2·08 | 74·9 2·03 | 97·3 0·545 | 61·1 0·264 |
| 6 | 71·2 1·85 | 92·7 3·03 | 17·6 1·54 | 32·7 0·795 | 51·7 2·52 | 74·2 2·34 | 800·2 0·596 | 52·9 0·341 |
| 7 | 65·9 2·16 | 90·8 3·47 | 19·0 1·74 | 24·9 0·974 | 47·3 2·96 | 73·3 2·64 | 02·7 0·639 | 44·3 0·411 |
| 8 | 60·0 2·46 | 88·3 3·93 | 19·8 1·96 | 16·7 1·161 | 42·6 3·41 | 72·0 2·94 | 04·8 0·676 | 35·2 0·472 |
| 9 | 53·7 2·76 | 85·2 4·39 | 19·8 2·18 | 08·0 1·357 | 37·4 3·86 | 70·3 3·23 | 06·5 0·706 | 25·5 0·525 |
| 10 | 46·8 3·04 | 81·8 4·86 | 19·8 2·40 | 699·1 1·56 | 31·9 4·33 | 68·2 3·51 | 07·9 0·730 | 15·4 0·572 |
| 11 | 39·5 3·33 | 77·7 5·33 | 19·3 2·63 | 89·8 1·77 | 26·0 4·80 | 65·6 3·79 | 08·8 0·750 | 04·7 0·611 |
| 12 | 31·6 3·60 | 73·2 5·81 | 18·1 2·87 | 80·2 1·99 | 19·8 5·27 | 62·7 4·07 | 09·3 0·765 | 593·5 0·644 |
| 13 | 23·3 3·87 | 68·2 6·29 | 16·2 3·10 | 69·7 2·23 | 12·9 5·75 | 59·4 4·33 | 09·4 0·777 | 81·8 0·672 |
| 14 | 14·5 4·14 | 62·7 6·77 | 14·2 3·35 | 60·0 2·44 | 05·7 6·23 | 55·7 4·60 | 09·1 0·784 | 69·6 0·694 |
| 15 | 05·1 4·40 | 56·7 7·26 | 11·6 3·59 | 48·1 2·67 | 698·3 6·73 | 51·6 4·86 | 08·4 0·788 | 56·9 0·712 |
| 16 | 595·2 4·65 | 50·2 7·76 | 08·4 3·84 | 36·8 2·90 | 90·3 7·22 | 47·0 5·11 | 07·3 0·791 | 43·7 0·726 |
| 17 | 84·9 4·91 | 43·2 8·25 | 04·7 4·09 | 24·9 3·14 | 81·8 7·71 | 42·1 5·37 | 05·7 0·791 | 30·1 0·736 |
| 18 | 74·0 5·16 | 35·7 8·75 | 00·6 4·34 | 12·7 3·39 | 73·2 8·21 | 36·7 5·62 | 03·7 0·789 | 15·9 0·743 |
| 19 | 62·8 5·41 | 27·7 9·25 | 695·9 4·59 | 00·1 3·63 | 63·9 8·71 | 30·9 5·87 | 01·3 0·785 | 01·2 0·749 |
| 20 | 51·0 5·66 | 19·3 9·75 | 90·8 4·85 | 587·1 3·88 | 54·1 9·21 | 24·6 6·12 | 798·5 0·780 | 486·1 0·750 |
| 21 | 38·7 5·91 | 10·3 10·25 | 85·2 5·11 | 73·6 4·13 | 44·1 9·72 | 18·0 6·37 | 95·2 0·773 | 70·5 0·750 |
| 22 | 25·9 6·15 | 00·8 10·76 | 79·1 5·36 | 59·7 4·38 | 33·6 10·22 | 10·9 6·61 | 91·5 0·767 | 54·3 0·748 |
| 23 | 12·7 6·40 | 590·9 11·26 | 72·5 5·62 | 45·4 4·63 | 22·7 10·73 | 03·3 6·86 | 87·4 0·759 | 37·7 0·746 |
| 24 | 498·9 6·64 | 80·5 11·77 | 64·5 5·88 | 30·7 4·89 | 11·3 11·23 | 695·4 7·10 | 82·8 0·750 | 20·7 0·740 |
| 25 | 6·88 | | 57·9 6·14 | | | 87·0 7·34 | 77·8 0·742 | |
| 26 | | | 49·9 6·39 | | | 78·2 7·59 | 72·3 0·733 | |
| 27 | | | 41·4 6·66 | | | 68·9 7·83 | 66·4 0·723 | |
| 28 | | | 32·4 6·91 | | | 59·1 8·07 | 60·1 0·713 | |
| 29 | | | 22·9 7·18 | | | 49·0 8·31 | 53·3 0·701 | |
| 30 | | | | | | 38·4 7·02 | | |

VOL
20
196

populated than the *F2*-levels of the doublet. If this were the case, one can use the Boltzmann equation, the transition probabilities at the $Q_1$ and $R_2$ heads, and a temperature of the order of 250°K to show that the observed $Q_1$ head would be more intense than the $R_2$ head by a factor of about 1·6. The observed profiles (Fig. 2) indicate an intensity ratio of only 1·1, from which it is necessary to conclude that equilibrium is not established in the $^2\Pi$ state.

The Meinel bands are probably excited from the ground state of the $N_2$ but for the purpose of simplifying the calculations we have assumed excitation from the ground state of $N_2^+$. This assumption will not seriously affect the results since the $B$ values for the $N_2^+$ and $N_2$ ground states are equal to within 4 per cent.

Working with these assumptions, the relative intensities of each of the lines were calculated according to

$$\text{intensity} \propto \nu^4 i(J',\lambda_{v'}) \exp\left\{-\frac{hc}{k}\frac{F''(K'')}{T_{\text{rot}}}\right\}$$

where $\nu$ is the frequency and $i(J',\lambda_{v'})$ is the line strength or intensity factor calculated from the expressions given by EARLS (1935) for $^2\Pi \rightarrow {}^2\Sigma$ transitions, and given in Table 1. The exponential term is the Boltzmann factor for the lower state. The factor representing the intensity alternation, not indicated in the above equation, was included in the calculations.

Another approximation is involved in the use of $K''$ in $F(K'')$ in the Boltzmann factor. For our purposes it is necessary to know the population distribution in the upper state due to excitation from the lower state when the mode of excitation is assumed to preserve the angular momentum. In the $^2\Sigma$ state, always strictly Hund's case (b), the angular momentum is characterized by the quantum number $K''$ which appears in the Boltzmann factor $\exp[-hcF''(K'')/kT]$. In the (upper) $^2\Pi$ state, however, for very small values of the rotational energy the state can be described as case (a) for which $J'$ but not $K'$ is defined; for larger energies ($K'' \sim 5$) the state is intermediate between cases (a) and (b) and neither $J'$ nor $K'$ is defined; and for still larger values where the state is case (b), $K'$ and $J'$ are defined. Consequently there is no simple way to determine the upper state population, given the lower, as in the case of the $^2\Sigma \rightarrow {}^2\Sigma$ bands. This problem could be overcome by the calculation of the upper state population with the aid of the intensity factors on the basis of an assumed excitation mechanism and lower state population. But because of the labour and uncertainties involved in such calculations, it was assumed that the resultant upper state population would be given to enough accuracy from the assumption that the population transfer went as $K'' = K'$. This assumption cannot be correct for all values of $K$ concerned but the error introduced will be trivial.

The calculated wavenumbers and intensities were used to construct model profiles by redistributing the line intensities according to the instrumental profile which was represented by a normal distribution with a standard deviation $\sigma = 4{\cdot}47$ cm$^{-1}$. The results of these calculations are shown in Fig. 2, where they are compared with the (2,0) and (3,1) band observations. The Fortrat diagram for the (3,1) band is practically identical with that of the (2,0) band, allowing a simple superposition and comparison of the two profiles on a wavenumber scale.



VOL. 20 1961

## 4. Conclusions

### 4.1. *Contamination*

It is clear from Fig. 2 that there is a considerable amount of contamination of the (2,0) band and before discussing the possible values of $T_{\rm rot}$ it is convenient to consider the nature of these contaminants. A comparison of the (2,0) and (3,1) band observations indicates the following regions of contamination: (1) the region around the low intensity ${}^{S}R_{21}$ head, (2) the minimum between the $R_2$ and $R_1$ heads, and (3) the intense and apparently red-degraded feature at $\tilde{\nu} = 12{,}634$ cm$^{-1}$ (7914 Å). The first two are no doubt due in part to the OH night-sky emission which is shown explicitly in Fig. 2. The analysis by Chamberlain and Roesler (1955) shows relatively strong OH (9,4) lines occurring at 12,781·1 cm$^{-1}$ and 12,730·0 cm$^{-1}$, and these are seen to fall in the first two regions of contamination mentioned above. Also, the intense feature at 12,634 cm$^{-1}$ may be partly explained by the strong OH (5,1) $Q$-branch head which is degraded to short wavenumbers.

Despite this coincidence of the $Q$ branch with the 12,634 cm$^{-1}$ feature, the identification is still uncertain for several reasons. First, there is the result obtained by Vallance Jones' (1960) analysis of the aurora of 10/11 February 1958 in which the feature at 12,634 cm$^{-1}$ appeared much too intense in comparison with other nearby OH lines and was then attributed, at least in part, to an unknown source. Further doubt as to the possible airglow source of these contaminations arises when it is seen from Fig. 2 that in all three observations of the (2,0) band and in all the forementioned regions of contamination, the intensities are the same relative to the three major $N_2^+$ peaks. This enhancement would imply that the contamination is auroral and not night-sky emission. Further attempts to identify the 12,634 cm$^{-1}$ feature lead only to the $NH_3$ head at 12,628 cm$^{-1}$ (Pearse and Gaydon, 1950), but this seems highly unlikely. It is hoped that with higher resolution on future plates, these identifications can be made more certain.

### 4.2. *Rotational temperature*

Regardless of the difficulty in identifying the possible contaminants, it is clear from the superposition of the observed (2,0) and (3,1) band profiles that certain regions of the (2,0) band are strongly enhanced over the (3,1) band, and consequently must be rejected in the determination of the rotational temperature because of contamination.

Fig. 2 indicates that the best fit with the observations requires a rotational temperature intermediate between 250°K and 500°K. The most temperature-sensitive regions appear to be the $R_1$ head, the minima on both sides of this head and the general slope of the profile on the short wavenumber side of the $Q_1$ head. All of these temperature-sensitive regions suffer relative increases or decreases in intensity as the temperature of the model increases or decreases respectively, and this general tendency continues for values of $T_{\rm rot}$ somewhat greater and less than those chosen here. It is clear that 500°K is too high: the $R_1$ head rises substantially above the observational envelope, and the surrounding minima are becoming too shallow. (The minimum between the $R_2$ and $R_1$ heads of the (2,0) band has been considerably contaminated; in this region comparisons should be made with the (3,1) band observation.) Finally, the parts of the profile at the short wavenumber



VOL
20
196

side of the $Q_1$ peak are generally too intense. (At wavenumbers short of 12,650 $cm^{-1}$, the model profile should be compared with the (3,1) band observation.) In all of these temperature-sensitive regions where the model for $T_{rot} = 500°K$ exceeds the observed intensities, it is seen that the $T_{rot} = 250°K$ model lies considerably below the observational envelope. The model profile for 375° agrees reasonably well with the observations, but may be slightly too hot for the region short of 12,650 $cm^{-1}$.

None of the models gives good agreement at the $R_2$ head except those for $T_{rot} < 250°K$. Since such a temperature would be completely unsatisfactory for the rest of the profile, the $R_2$ head was given little weight in these considerations. The small observed maximum at $\tilde{\nu} = 12{,}798$ $cm^{-1}$ cannot be completely explained by the ${}^SR_{21}$ head, which occurs at a significantly larger wavenumber.

## 5. Summary

Of the models calculated, the one for $T_{rot} = 375°K$ provides the best fit with the observations. Of course, no great precision can be ascribed to this value, and in view of the uncertainties involved it appears from Fig. 2 that conservative upper and lower limits to the rotational temperature are 425°K and 275°K with the best value being about 350°K. Approximate models were calculated at other values of $T_{rot}$, and they are entirely consistent with the general interpretation given here.

A rotational temperature of $350 \pm 75°K$ for the $A \rightarrow X$ band is comparable with the higher determinations from the $B \rightarrow X$ bands. The latter are $330 \pm 40°K$ (Vallance Jones *et al.*, 1953), a range from 251 to 376°K (Vallance Jones and Harrison, 1955), and a variation from 150 to 465°K with a mean of 258°K (Shepherd and Hunten, 1955). Wallace (1959) obtained a variation of $225°K \leqslant T_{rot} \leqslant 575°K$ but this was from an exceptionally high-altitude aurora whereas the $A \rightarrow X$ band observations are not.

About all that can be said of this determination is that it is not necessarily at variance with the $A \rightarrow X$ temperatures. A temperature accurate enough for comparison with the $A \rightarrow X$ bands would require an approximately fourfold increase in resolution over that used here. Such a well-exposed spectrum might be obtained of an aurora as bright as that of 11 February 1958, but for the purposes of the suggested comparison, it would be necessary to obtain a spectrum of the $A \rightarrow X$ bands at the same time and of the same part of the sky.

*Acknowledgement*—It is a pleasure for one of us (W. G. M.) to express his appreciation to the National Science Foundation for a fellowship which made this work possible.

## References


| | | |
|---|---|---|
| Chamberlain J. W. and Roesler F. L. | 1955 | *Astrophys. J.* **121,** 541. |
| Coster D. and Brons H. H. | 1931 | *Z. Phys.* **73,** 747. |
| Dalby F. W. and Douglas A. E. | 1951 | *Phys. Rev.* **84,** 843. |
| Douglas A. E. | 1953 | *Astrophys. J.* **117,** 380. |
| Earls L. T. | 1935 | *Phys. Rev.* **48,** 423. |
| Herzberg G. | 1950 | *Spectra of Diatomic Molecules* (2nd ed.). Van Nostrand, New York. |
| Lui I. D. | 1959 | *Astrophys. J.* **129,** 516. |




VOL. 20 1961

| | | |
|---|---|---|
| Meinel A. B. | 1951 | *Astrophys. J.* **114,** 431. |
| Pearse R. W. B. and Gaydon A. G. | 1950 | *The Identification of Molecular Spectra* (2nd ed. rev.). Chapman and Hall, London. |
| Roesler F. L., Fan C. Y. and Chamberlain J. W. | 1958 | *J. Atmosph. Terr. Phys.* **12,** 200. |
| Shepherd G. G. and Hunten D. M. | 1955 | *J. Atmosph. Terr. Phys.* **6,** 328. |
| Vallance Jones A. | 1960 | Private communication. |
| Vallance Jones A. and Harrison A. W. | 1955 | *J. Atmosph. Terr. Phys.* **6,** 336. |
| Vallance Jones A., Hunten D. M. and Shepherd G. G. | 1953 | *Astrophys. J.* **118,** 350. |
| Wallace L. | 1959 | *J. Atmosph. Terr. Phys.* **17,** 46. |


VOL.
20
1961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 9 to 18. Pergamon Press Ltd. Printed in Northern Ireland

# The scattering of radio waves by an extended randomly refracting medium*

S. A. Bowhill
Ionosphere Research Laboratory, The Pennsylvania State University

(*Received* 19 *July* 1960)

**Abstract**—A continuous medium containing three-dimensional random inhomogeneities of refractive index scatters an electromagnetic wave which is incident upon it. This paper derives the form of the emerging angular power spectrum of the wave, for the case where the scales of the inhomogeneities are different in the three spaced directions.

It is shown that the medium cannot be analysed as a series of superposed thin phase screens, spaced in the propagation direction, and with independent phase profiles. When a limiting process is carried out, allowing for an arbitrary spatial correlation function for the inhomogeneities in the three directions, a quite different result is obtained. This disparity is due to diffractive changes in the wave in passing from one inhomogeneity to the next.

## 1. Introduction

It has been known for a long time that the ionosphere does not reflect radio waves like a smooth mirror. Some random variation in its properties must occur, to explain the observed reflected wave. Various lines of physical reasoning indicate that irregular spatial and temporal electron density variations must be present. The magnitude of these variations, their horizontal and vertical scales, and their rate of change with time, are usually taken as functions of height only, and not of time, or of horizontal position—this is equivalent to postulating the statistical stationarity of the electron density variation.

Very little knowledge is available about the variation with height of these four statistical measures. The only useful data are those deriving from radio reflection measurements; they are, however, subject to the following difficulties:

(i) It is not easy to determine experimentally all the properties of a randomly reflected signal arriving at the ground.

(ii) The signal observed at the ground is not identical, even as regards its statistical measures, with the signal which emerged from the ionosphere, due to diffractive changes.

(iii) Any single radio frequency, being reflected within the ionosphere, is acted upon by irregularities of electron density at all heights, up to and including the height of reflection.

The first of these problems can be overcome by suitable experimental planning: the second has been effectively solved by the work of Booker *et al.* (1950), Hewish (1951, 1952) and Ratcliffe (1956). The third, however, is still in the early stages of solution; for instance, Pitteway (1958) has shown how the signal scattered by an ionosphere may be calculated for a *particular* arrangement of weak irregularities.

One approach to this problem is to consider the diffraction of a wave by a thick

* The research described in this paper has been supported by the Geophysics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command, under Contract AF(19)604-1304.

slab containing inhomogeneities of refractive index. Fejer (1953) examined a model consisting of a number of thin shallow phase-modulating diffracting screens with the same known autocorrelogram of phase variation across each of them. They were assumed to be sufficiently far apart in the mean direction of propagation for the phase variation to be dissimilar at adjacent screens. The power spectrum of the wave emerging from a particular screen, therefore, was obtained by convoluting the incident angular power spectrum with the spatial power spectrum of the screen.

This work will be extended in Section 3 to the case of an inhomogeneous refracting medium with a continuous variation of its properties in the direction of the incident wave: it will be shown that under certain circumstances this can lead to an angular spectrum of scattered waves of a different form from that given by Fejer. This difference is due to the limiting process used in the analysis.

## 2. Discussion of the Work of Hewish, Fejer and Bramley

The work of Fejer (1953) provided a connecting link between the theories of weak and strong scatter. A scattering process is said to be weak when the original incident wave emerges from the medium substantially unchanged. It is then known as the "specular" component, by analogy with specular (as opposed to diffuse) reflection in optics. The interaction of the scattered field with the irregularities (producing what may be termed second order scatter) is negligible under these circumstances. This corresponds to the Born approximation in atomic theory. The amplitude of the signal a long way from a weak scattering screen of this kind has a "displaced Gaussian" probability distribution in the sense used by McNicol (1949).

When a number of weakly scattering screens are superposed, the scattering power of the combination is no longer weak, and the specular component may become completely scattered away. Under this condition of strong scattering, the mechanism of propagation of the wave in its latter stages consists of cross-scattering, by new irregularities, of waves which have already been scattered previously. These waves give a signal a long way away from the medium which has an amplitude probability distribution of the Rayleigh form. It must be emphasized that the presence of multiple scattering does not necessarily imply that the wave bundle emerging from the medium has a very wide cone angle (i.e., that the layer is completely irregular). An example of this type of propagation is the so-called "forward scatter" mode of v.h.f. propagation. This aspect of ionospheric scattering will not be considered here.

Bramley (1954) showed that a result identical to Fejer's could be obtained by assuming the whole medium to be "collapsed" in the direction of the incident wave. It then becomes a single thin deeply modulated phase screen of the kind treated by Hewish (1951, 1952). This artifice is equivalent to calculating the phase of the wave emerging from the medium by integrating the phase along a line projected back parallel to the incident direction, a result which was also found by Feinstein (1954). It will be shown in Section 3 that this is not true in general, even for the case of a weakly scattering medium of finite extent.

The discussion so far has dealt exclusively with scattering due to changes in

refractive index. In an ionospheric layer, however, it is possible for random inhomogeneities in absorption to occur, as was shown by JONES *et al.* (1953). For a weakly scattering medium, such as that discussed in Section 3, the equivalence between shallow phase and amplitude screens (BOWHILL, 1957) enables the amplitude case to be evaluated immediately. The case of strong scatter by amplitude inhomogeneities is not of importance practically, as the emerging scattered wave would necessarily be very highly attenuated.

## 3. WEAK SCATTER WITH THREE-DIMENSIONAL PARTIAL COHERENCE

A statistically homogeneous ionized medium, of thickness $t$, has a plane wave incident on it normally from above (Fig. 1). Inhomogeneities are present in the medium which alter the amplitude and phase of the wave passing through it. The

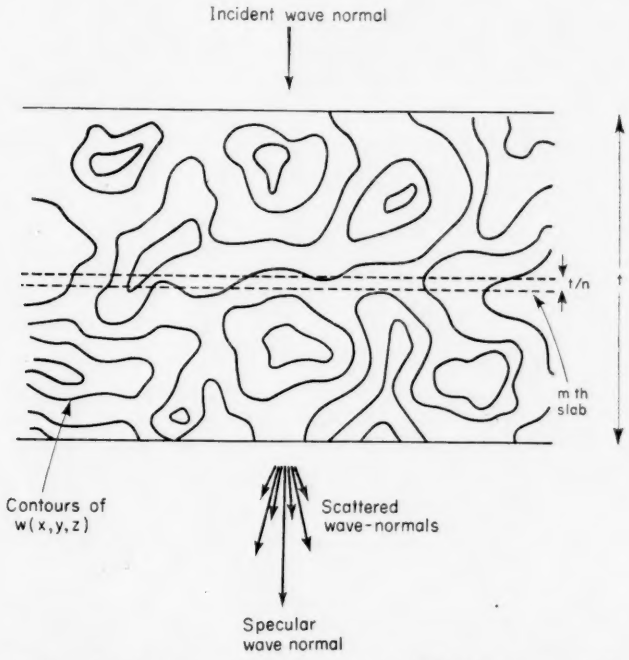


Fig. 1. The scattering of a plane wave by a three-dimensional inhomogeneous refracting medium.

angular power spectrum of the wave emerging from the medium will be calculated. As mentioned above, the assumption is made that the incident wave passes through the medium substantially unchanged.

The method of treatment is to split up the medium into a number $n$ of slabs of equal thickness $t/n$, lying parallel to $(x, y)$ plane. Each of these slabs is treated as a thin two-dimensional diffracting screen, and the wave emerging from the medium is evaluated. The continuous medium is then taken as the limit of this assembly of screens as $n \to \infty$. The pattern of the irregularities in adjacent slabs is highly correlated; this is an essential feature of the analysis. In contrast, Fejer's



VOL. 20 1961

treatment assumes complete independence of adjacent slabs, and so cannot introduce the concept of correlation in the $z$ direction.

For convenience in following the analysis, it is split up into a number of steps.

3.1. *Effect of the* m*th slab on a single plane wave*

A plane wave incident normally on a slab continues through it with negligible change in amplitude, accompanied by a scattered wave. Let the phase refractive index at any point in the slab be

$$\mu = 1 + w(x, y, z) \tag{1}$$

where $w(x, y, z)$ represents the deviation of the refractive index from unity, and $z = mt/n$. The mean value of $w(x, y, z)$ will be taken as zero, to simplify the argument: this does not affect the generality of the conclusions. The phase path through the slab is

$$\frac{\mu t}{n} = \frac{(1 + w)t}{n}$$

and so the deviation in the phase from the mean is

$$\left(\frac{2\pi t}{\lambda n}\right) w(x, y, z) \tag{2}$$

where $\lambda$ is the wavelength of the radiation. The complex electric field at the wave immediately after the passing through the screen is therefore

$$\underline{E} = 1 + \left(\frac{2\pi i t}{\lambda n}\right) w(x, y, z)$$

and the emerging wave consists of a specular component together with a scattered wave, the angular spectrum of which is

$$p_m(s, r) = \frac{2\pi i t}{\lambda n} \int\int_{-\infty}^{\infty} w(x, y, z) \exp\left[\frac{-2\pi i(sx + ry)}{\lambda}\right] dx\, dy. \tag{3}$$

3.2. *Formation of spectra at successive slabs*

In calculating the effect of the next slab on the wave emerging from the first slab, it is assumed that both slabs can be regarded as ideal infinitely thin phase screens, separated by a distance $t/n$. This is justifiable, since each slab is very thin. In propagating the distance $t/n$, the scattered components of the angular spectrum change in phase relative to the specular component by an amount

$$\frac{\pi i t(r^2 + s^2)}{n\lambda}$$

provided $r$ and $s$ are small. The signal emerging from the second screen therefore consists of four parts:

(a) the specular component;



VOL
20
196

(b) that part of the specular component which has been scattered by the second screen: namely, $p_2(s, r)$;

(c) the signal $p_1(s, r) \exp [\pi it(r^2 + s^2)/n\lambda]$ scattered by the first screen, which passes through the second screen unaffected, as though it were a specular signal;

(d) a very small signal which has been scattered by both screens. This will be neglected, since only weak scatter is being considered.

The spectrum $P_2(s, r)$ of the scattered signal emerging from the second screen is therefore

$$P_2(s, r) = p_2(s, r) + p_1(s, r) \exp\left[\frac{\pi it(r^2 + s^2)}{n\lambda}\right].$$

Similarly

$$P_3(s, r) = p_3(s, r) + p_2(s, r) \exp\left[\frac{\pi it(r^2 + s^2)}{n\lambda}\right] + p_1(s, r) \exp\left[\frac{2\pi it(r^2 + s^2)}{n\lambda}\right]$$

up to

$$P_n(s, r) = p_n(s, r) + p_{n-1}(s, r) \exp\left[\frac{\pi it(r^2 + s^2)}{n\lambda}\right] + \ldots . + p_1(s, r) \exp\left[\frac{(n-1)\pi it(r^2 + s^2)}{n\lambda}\right]. \quad (4)$$

3.3. *The emerging angular power spectrum*

The angular power spectrum $W_n(s, r)$ of the scattered wave emerging from the scattering medium is given by

$$W_n(s, r) = \overline{P_n(s, r) P_n^*(s, r)} \quad (5)$$

the average being taken over systems. Putting, for convenience,

$$u = \exp [\pi it(r^2 + s^2)/n\lambda] \quad (6)$$

in (4), and noting that $u^* = u^{-1}$,

$$\left.\begin{aligned} P_n(s, r) &= p_n(s, r) + up_{n-1}(s, r) + \cdots + u^{n-2}p_2(s, r) + u^{n-1}p_1(s, r) \\ \text{and} \\ P_n^*(s, r) &= p_n^*(s, r) + u^{-1}p_{n-1}^*(s, r) + \cdots + u^{-n+2}p_2^*(s, r) + u^{-n+1}p_1^*(s, r) \end{aligned}\right\}. \quad (7)$$

Substituting from (7) into (5),

$$\begin{aligned} W_n(s, r) = {} & \overline{p_n(s, r)p_n^*(s, r)} + \overline{p_{n-1}(s, r)p_{n-1}^*(s, r)} + \cdots + \overline{p_1(s, r)p_1^*(s, r)} + \\ & + u\,\overline{p_{n-1}(s, r)p_n^*(s, r)} + \overline{p_{n-2}(s, r)p_{n-1}^*(s, r)} + \cdots + \overline{p_1(s, r)p_2^*(s, r)} + \\ & + u^{-1}[p_n(s, r)p_{n-1}^*(s, r) + p_{n-1}(s, r)p_{n-2}^*(s, r) + \cdots + p_2(s, r)p_1^*(s, r) + \cdots] + \\ & + u^k \sum_{m=1}^{n-k} \overline{p_m(s, r)p_{m+k}^*(s, r)} + u^{-k} \sum_{m=1}^{n-k} \overline{p_m^*(s, r)p_{m+k}(s, r)} + \ldots\ldots\ldots\ldots \end{aligned} \quad (8)$$

The coefficients of $u^k$ above are all sums of terms of the form $\overline{p_m(s, r)p_{m+k}^*(s, r)}$. This quantity is evaluated in the next section.



VOL. 20 1961

3.4. *Evaluation of* $\overline{p_m(s, r)p^*_{m+k}(s, r)}$

From (3),

$$p_m(s, r) = \frac{2\pi it}{\lambda n}\int\int_{-\infty}^{\infty} w\left(x, y, \frac{mt}{n}\right)\exp\left[\frac{-2\pi i(sx + ry)}{\lambda}\right]dx\,dy$$

$$p^*_{m+k}(s, r) = \frac{-2\pi it}{\lambda n}\int\int_{-\infty}^{\infty} w\left[x', y', \frac{(m + k)t}{n}\right]\exp\left[\frac{2\pi i(sx' + ry')}{\lambda}\right]dx'\,dy'$$

where $x'$ and $y'$ are variables similar to $x$ and $y$. So

$$P_m(s, r)p^*_{m+k}(s, r) = \left(\frac{2\pi t}{\lambda n}\right)^2\int\int\int\int_{-\infty}^{\infty} w\left(x, y, \frac{mt}{n}\right)w\left[x', y', \frac{(m + k)t}{n}\right]\times$$
$$\times\exp\left\{\frac{2\pi i[s(x' - x) + r(y' - y)]}{\lambda}\right\}dx\,dy\,dx'\,dy'.$$

Putting $x' = x + \alpha$, $y' = y + \beta$, this becomes

$$p_m(s, r)p^*_{m+k}(s, r) = \left(\frac{2\pi t}{\lambda n}\right)^2\int\int\int\int_{-\infty}^{\infty} w\left(x, y, \frac{mt}{n}\right)w\left[x + \alpha, y + \beta, \frac{(m + k)t}{n}\right]\times$$
$$\times\exp\left[\frac{2\pi i(s\alpha + r\beta)}{\lambda}\right]dx\,dy\,d\alpha\,d\beta.$$

Averaging over systems, and integrating with respect to $x$ and $y$,

$$\overline{p_m(s, r)p^*_{m+k}(s, r)} = \left(\frac{2\pi t}{\lambda n}\right)^2 S.\quad\int\int_{-\infty}^{\infty}\overline{w\left(x, y, \frac{mt}{n}\right)w\left[x + \alpha, y + \beta, \frac{(m + k)t}{n}\right]}\times$$
$$\times\exp\left[\frac{2\pi i(s\alpha + r\beta)}{\lambda}\right]d\alpha\,d\beta$$

where $S$ is the area of the slab in the $(x, y)$ plane. Now it will be assumed that $w(x, y, z)$ is a stationary random function of the three variates $x$, $y$ and $z$, with zero mean. Under these circumstances,

$$w\left(x, y, \frac{mt}{n}\right)w\left[x + \alpha, y + \beta, \frac{(m + k)t}{n}\right] = w^2(x, y, z)\rho\left(\alpha, \beta, \frac{kt}{n}\right) \tag{9}$$

where $\rho(\alpha, \beta, \gamma)$ is the correlogram of $w(x, y, z)$ in three dimensions. So

$$\overline{p_m(s, r)p^*_{m+k}(s,r)} = \left(\frac{2\pi t}{\lambda n}\right)^2 S\,\overline{w^2}\int_{-\infty}^{\infty}\rho\left(\alpha, \beta, \frac{kt}{n}\right)\exp\left[\frac{2\pi i(s\alpha + r\beta)}{\lambda}\right]d\alpha\,d\beta. \tag{10}$$

Since the correlogram must be an even function of $\alpha$, $\beta$ and $\gamma$, it follows that

$$\overline{p_m(s, r)p^*_{m+k}(s, r)} = \overline{p_m{}^*(s, r)p_{m+k}(s, r)}. \tag{11}$$

In addition, it is evident that these quantities are independent of $m$. It follows from this that all the terms in the coefficient of $u^k$ in (8) must be equal.

3.5. *General form for the emerging power spectrum*

Rewriting (8) in summation form,

$$W_n(s, r) = \sum_{m=1}^{n} \overline{p_m(s, r)p_m{}^*(s, r)} + \sum_{k=1}^{n-1} u^k \sum_{m=1}^{n-k} \overline{p_m(s, r)p^*_{m+k}(s, r)} + \\ + \sum_{k=1}^{n-1} u^{-k} \sum_{m=1}^{n-k} \overline{p^*_m(s,r)p_{m+k}(s,r)}$$

Substituting from (10),

$$W_n(s, r) = \left(\frac{2\pi t}{\lambda n}\right)^2 S\,\overline{w^2} \iint_{-\infty}^{\infty} \exp\left[\frac{2\pi i(s\alpha + r\beta)}{\lambda}\right] \times \\ \times \left\{n\rho(\alpha, \beta, 0) + 2\sum_{k=1}^{n=1} (n - k)\rho\left(\alpha, \beta, \frac{kt}{n}\right) \cos\left[\frac{k\pi t(r^2 + s^2)}{n\lambda}\right]\right\} dx\, d\beta. \quad (12)$$

The thickness of the individual slabs will now be taken as very small, and the quantity in curly brackets will be replaced by the corresponding integral

$$2\int_0^n (n - k)\rho\left(\alpha, \beta, \frac{kt}{n}\right) \cos\left[\frac{k\pi t(r^2 + s^2)}{n\lambda}\right] dk + 0\left(\frac{1}{n}\right).$$

Putting $kt/n = \gamma$ (representing incremental distance in the $z$ direction), this becomes

$$\left(\frac{2n^2}{t^2}\right) \int_0^t (t - \gamma)\rho(\alpha, \beta, \gamma) \cos\left[\frac{\pi\gamma(r^2 + s^2)}{\lambda}\right] d\gamma + 0\left(\frac{1}{n}\right).$$

Substituting into (12), and taking the limit $n \to \infty$,

$$W_n(s, r) = \left(\frac{8\pi S\overline{w^2}}{\lambda^2}\right) \int\int_0^{\infty} d\alpha\, d\beta \int_0^t (t - \gamma)\rho(\alpha, \beta, \gamma) \cos\left[\frac{\pi\gamma(r^2 + s^2)}{\lambda}\right] d\gamma. \quad (13)$$

3.6. *Case of a slab of medium*

Equation (13) cannot be carried further in its general form for a slab of thickness $t$, comparable with the average dimensions of the irregularities in the $z$ direction. However, the case of most interest (for comparison with Fejer's results) concerns a slab which has several layers of irregularities in its depth. Expressed mathematically, this means that $\rho(\alpha, \beta, \gamma) \sim 0$ for all $\gamma > d$, where $d$ is some parameter $\ll t$. Because of this, the integral in (13) will have a very small contribution from values of $\gamma > d \ll t$. This means, firstly, that the integral can be taken over the range $0 < \gamma < \infty$, and, secondly, that the term $(t - \gamma)$ can be replaced by $t$. So

$$W_n(s,r) = \left(\frac{8\pi^2 S\overline{w^2}t}{\lambda}\right) \int\int_{-\infty}^{\infty} \exp\left[\frac{2\pi i(s\alpha + r\beta)}{\lambda}\right] d\alpha\, d\beta \times \int_0^{\infty} \rho(\alpha,\beta,\gamma) \cos\left[\frac{\pi\gamma(r^2 + s^2)}{\lambda}\right] d\gamma$$

$$= \left(\frac{4\pi^2 S\overline{w^2}}{\lambda^2}\right) \iiint_{-\infty}^{\infty} \rho(\alpha, \beta, \gamma) \exp\{2\pi i[s\alpha + r\beta + 2(r^2 + s^2)]/\lambda\}\, d\alpha\, d\beta\, d\gamma.$$



VOL. 20 1961

This integral should be compared with the three-dimensional spatial power spectrum $V(\nu, \chi, \omega)$ of $w(x, y, z)$, which is given by

$$V(\nu, \chi, \omega) = \int\int\int_{-\infty}^{\infty} \rho(\alpha, \beta, \gamma) \exp\left[\frac{2\pi i(\nu\alpha + \chi\beta + \omega\gamma)}{\lambda}\right] d\alpha\, d\beta\, d\gamma. \tag{14}$$

It should be noted that $V(\nu, \chi, \omega)$ is the spatial power spectrum *per unit volume* of the medium. The emerging power spectrum therefore, can be written

$$W_n(s, r) = \left(\frac{4\pi^2 S\overline{w^2}t}{\lambda^2}\right) V\left[s, r, \tfrac{1}{2}(r^2 + s^2)\right]. \tag{15}$$

3.7. *Medium with Gaussian spatial correlogram*

As an illustration of the way (14) and (15) can be applied, it will be assumed that

$$\rho(\alpha, \beta, \gamma) = \exp\left(-\frac{\alpha^2}{2{D_1}^2} - \frac{\beta^2}{2{D_2}^2} - \frac{\gamma^2}{2{D_3}^2}\right). \tag{16}$$

$D_1$, $D_2$ and $D_3$ are the structure sizes of the irregularities in the $x$, $y$ and $z$ directions, respectively. Taking the Fourier transform of (16), as given by (14),

$$V(\nu, \chi, \omega) = (2\pi)^{3/2} D_1 D_2 D_3 \times \exp\left[\frac{-2\pi^2({D_1}^2\nu^2 + {D_2}^2\chi^2 + {D_3}^2\omega^2)}{\lambda^2}\right]$$

and substituting into (15)

$$W_n(s, r) = (2\pi)^{7/2}\left(\frac{S\overline{w^2}tD_1D_2D_3}{\lambda^2}\right) \times \exp\left\{-2\pi^2\left[{D_1}^2s^2 + {D_2}^2r^2 + \frac{{D_3}^2(r^2 + s^2)^2}{4}\right]/\lambda^2\right\}. \tag{17}$$

Equation (17) has several interesting features:

A. It is evident that the scattered power is proportional to the thickness of medium traversed, to the area of the medium in the $(x, y)$ plane, and to the mean square value of $w(x, y, z)$, all of which might be expected physically.

B. If $D_3$, the structure size of the irregularities of $w(x, y, z)$ in the $z$ direction, is much less than $D_1$ or $D_2$, the power spectrum becomes effectively

$$W_n(s, r) \propto \exp\left[\frac{-2\pi^2({D_1}^2s^2 + {D_2}^2r^2)}{\lambda^2}\right] \tag{18}$$

which has the same form as for the angular power spectrum of waves scattered from a thin screen with the same form of spatial correlogram in the $(x, y)$ plane.

C. If, on the other hand, the structure sizes $D_1$ and $D_2$ in the $x$ and $y$ directions are much *smaller* than the $z$ direction size $D_3$, the angular power spectrum does not become very broad, as it would in the case of scattering by a thin screen with $D_1$ and $D_2$ small. The form of the angular power spectrum becomes quite different:

$$W_n(s, r) \propto \exp\left[\frac{-2\pi^2{D_3}^2(r^2 + s^2)^2}{4\lambda^2}\right]. \tag{19}$$

The exponential part of the angular spectrum $W_n(s, r)$ of (17) is plotted as a function

of $s$ in Fig. 2 for $D_1 = D_2 = \lambda$, and for $D_3 = \lambda$ and $4\lambda$. For $D_3 = \lambda$, the angular power spectrum of the scattered wave is practically Gaussian in form. However, when $D_3 = 4\lambda$, corresponding to "stretching out" the irregularities in the $z$ direction, the spectrum has the form of (19), except near the origin of $s$.

Since $D_1 = D_2 = \lambda$ in both cases considered, the phase variation with $x$ and $y$ at the face of the slab, obtained by integrating through the medium along a path parallel to the $z$ direction, will have a spatial correlogram of Gaussian form in both cases. This, therefore, illustrates the fact mentioned in Section 2, that a thick slab containing irregularities of refractive index cannot be "collapsed" in the $z$ direction into an equivalent thin phase screen, even when the phase variation is much less than a radian, and the scatter is weak.

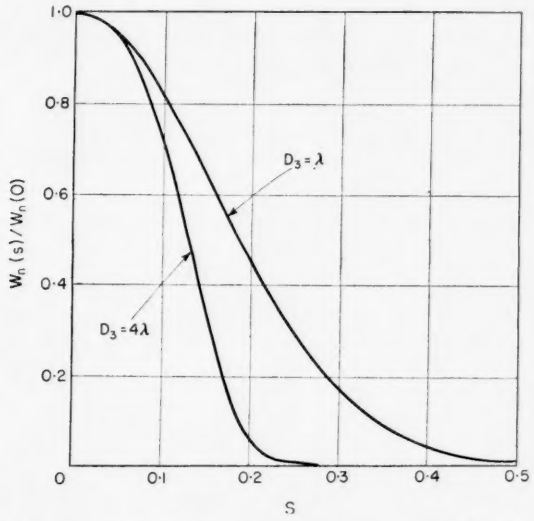


Fig. 2. The angular power spectrum emerging from a scattering slab, for $D_1 = D_2 = \lambda$, and two values of $D_3$.

## 4. Extension to the Case of Strong Scatter

Equation (17), which gives the angular spectrum for weak scatter by ionospheric irregularities, may be extended to the case of strong scatter in exactly the same way as (24) of Fejer's paper, to which it corresponds identically (with the addition of the diffractive term in the exponent, due to the finite value of $D_3$). To do this it is only necessary to choose a suitable thickness of the medium, such that (i) the scatter is weak, (ii) $t \gg D_3$, and then to combine a number of these slabs to give a thick, strongly scattering medium. It will be remembered that the reason for (17) being different from Fejer's result was that he neglected correlation effects between adjacent screens. Due to condition (ii) however, this effect is not important in the case considered here, because the correlation only occurs near the *boundaries* of adjacent slabs, and has a negligible effect.

It should be mentioned that if conditions (i) and (ii) cannot be satisfied simultaneously for a given medium, the scatter must be strong for one individual

irregularity, and the problem cannot then be solved by any known stochastic method.

The procedure Fejer used was to evaluate the power spectrum emerging from a slab by convoluting the *power* spectrum entering it with the spatial power spectrum of the slab itself. A simplification occurred in his analysis, in that the multiple convolution of any number of Gaussian functions always gives a function which is Gaussian in form. The presence of a term in $s^4$, in the exponent in (17) complicates the analytical solution of the problem, which would probably need to be solved numerically.

## 5. Summary

The central argument of this paper is that a continuous randomly inhomogeneous medium cannot be represented by a number of thin spaced random screens with uncorrelated phase profiles. By considering a continuous medium as the limit of a series of thin spaced screens, it is shown that the scale of the irregularities in the direction of propagation has an important influence on the angular power spectrum of the emerging wave.

## References


| | | |
|---|---|---|
| Booker H. G., Ratcliffe J. A. and Shinn D. H. | 1950 | *Phil. Trans.* A **242,** 579. |
| Bowhill S. A. | 1957 | *J. Atmosph. Terr. Phys.* **11,** 91. |
| Bramley E. N. | 1954 | *Proc. Roy. Soc.* A **225,** 515. |
| Feinstein J. | 1954 | *Trans. Inst. Radio Engrs* **AP-2,** 23, 63. |
| Fejer J. A. | 1953 | *Proc. Roy. Soc.* A **220,** 445. |
| Hewish A. | 1951 | *Proc. Roy. Soc.* A **209,** 81. |
| Hewish A. | 1952 | *Proc. Roy. Soc.* A **214,** 494. |
| Jones R. E., Millman G. H. and Nertney R. J. | 1953 | *J. Atmosph. Terr. Phys.* **3,** 79. |
| McNicol R. W. E. | 1949 | *Proc. Inst. Elect. Engrs* III **96,** 517. |
| Pitteway M. L. V. | 1958 | *Proc. Roy. Soc.* A **246,** 556. |
| Ratcliffe J. A. | 1956 | *Rep. Progr. Phys.* **19,** 188. |




VOL. 20 1961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 19 to 30. Pergamon Press Ltd. Printed in Northern Ireland

# The effective recombination coefficient of an ionosphere containing a mixture of ions*

S. A. Bowhill
Ionosphere Research Laboratory, The Pennsylvania State University
University Park, Pennsylvania

(*Received* 2 *August* 1960)

**Abstract**—The effective recombination coefficient of the normal ionosphere may be evaluated from the nocturnal variation of ionization density; from the rate of change of ionization during a solar eclipse; or from the "sluggishness" of the ionization, as exhibited in its diurnal variation. These three methods are found to give widely different results, at any one height in the ionosphere.

It is shown that an ionosphere containing a mixture of several ion species with different recombination coefficients, gives adequate explanation of these discrepancies. Formulae are presented for the effective recombination coefficient for this model.

A consideration of existing knowledge of ionization processes in the $E$-region suggests that the ions concerned may be those of molecular oxygen and nitric oxide; the latter being produced by ion–atom interchange between nitrogen molecules and atomic oxygen ions. The nocturnal $E$-region should then consist primarily of $NO^+$ ions.

## 1. Introduction

The determination of the effective recombination coefficient of the ionization in the ionosphere has been the subject of a number of papers in recent years. Good summaries have been given by Ratcliffe (1956, 1960). Because of the lack of knowledge of the detailed processes involved in the recombination, it has become commonplace to write the equation

$$\frac{dn_e}{dt} = q_e - \alpha n_e^2 \tag{1}$$

where $n_e$ = electron density at a given height at a time $t$;
$q_e$ = rate of ionization production at that height;
$\alpha$ = effective recombination coefficient.

Experimental values of electron density, and theoretical values of the electron production, are inserted into (1), and values of the effective recombination coefficient are deduced.

It has long been realized that widely different values of $\alpha$ may be obtained for the same ionospheric region, depending on the particular observations which are used in (1). For instance, in the $E$-region, values deduced from the nocturnal decay of ionization, observations during solar eclipses, and measurements of the so-called "sluggishness" of the layer (the time delay of the maximum ionization of the layer relative to noon), give values of $\alpha$ varying from $4 \times 10^{-8}$ to $10^{-9}$ cm$^3$ sec$^{-1}$. Various explanations have been brought forward for this disparity; possible non-uniformity of the ionizing radiation over the solar disk, the presence of negative ions (Mitra, 1957b), vertical motions of the layer (Appleton and Lyon, 1957), with only partial success.

* The research described in this paper was supported by the Guidance and Control Laboratory of the Army Ballistic Missile Agency, under Contract No. DA-36-061-ORD-577.

Bates and McDowell (1957) considered the possibility of explaining some features of the eclipse results on the basis of the presence of two positive ions. They based this on the fact that several constituents in any given ionized region are bound to be ionized, and that several species of ion will therefore certainly be present. Later work (Bates and Nicolet, 1960) has confirmed this supposition. The two-ion hypothesis has been criticized by Minnis (1958b) and further analysed by McElhinny (1959).

The purpose of this paper is to investigate in detail the form of the recombination when several species of positive ion are present, each recombining by a simple process of the form

$$X^+ + e \rightarrow X + h\nu$$
$$XY^+ + e \rightarrow X + Y$$

It will be shown that the use of (1) gives values for the effective recombination coefficient which depend critically on the type of experimental data used.

## 2. Nocturnal Recombination

The simplest case is that of nocturnal recombination, in which it will be assumed that $q_e = 0$. If $m$ species of positive ion are present, the $r$th species having a density $n_r$ and recombination coefficient $\alpha_r$, then at a given height

$$\frac{dn_r}{dt} = -\alpha_r n_r n_e. \tag{2}$$

Neglecting negative ion formation, for electrical neutrality

$$n_e = \sum_{r=1}^{m} n_r \tag{3}$$

and from (2)

$$\frac{dn_e}{dt} = -n_e \sum_{r=1}^{m} \alpha_r n_r. \tag{4}$$

For $q_e = 0$, the effective nocturnal recombination coefficient $\alpha_a$ is given from (1) by

$$\alpha_a = -\frac{1}{n_e^2}\frac{dn_e}{dt} = \sum_{r=1}^{m} (n_r/n_e)\alpha_r. \tag{5}$$

The effective recombination coefficient is therefore the arithmetic mean for all ions present, weighted by their relative abundance.

Equations (2) and (5) may be written in the form

$$\frac{d(\ln n_r)}{dt} = -\alpha_r n_e$$

$$\frac{d(\ln n_e)}{dt} = -\alpha_a n_e$$

whence

$$\frac{d[\ln (n_r/n_e)]}{dt} = (\alpha_a - \alpha_r)n_e. \tag{6}$$



VOL. 20 1961

Since $(n_r/n_e)$ is the relative abundance of the $r$th ion, it follows that all ions for which $\alpha_r > \alpha_a$ are decreasing, and ions for which $\alpha_r < \alpha_a$ are increasing, in relative abundance. The value of $\alpha_a$ must therefore decrease continuously with time. Its rate of increase may be found by rewriting (6) in the form

$$\frac{d(\alpha_r n_r/n_e)}{dt} = \alpha_r n_r(\alpha_a - \alpha_r). \tag{7}$$

Summing over $r$ and using (5),

$$\sum_{r=1}^{m} \frac{d(\alpha_r n_r/n_e)}{dt} = \alpha_a \sum_{r=1}^{m} \alpha_r n_r - \sum_{r=1}^{m} \alpha_r^2 n_r$$

$$\frac{d\alpha_a}{dt} = \alpha_a^2 n_e - \sum_{r=1}^{m} \alpha_r^2 n_r.$$

Defining $\alpha_b$ as the weighted r.m.s. value of $\alpha_r$,

$$\frac{d\alpha_a}{dt} = -n_e(\alpha_b^2 - \alpha_a^2) \tag{8}$$

where

$$\alpha_b^2 = \sum_{r=1}^{m} (n_r/n_e)\alpha_r^2. \tag{9}$$

Except for the degenerate case where all the $\alpha_r$ are equal, $\alpha_b > \alpha_a$ for all $\alpha_r$, and $\alpha_a$ is again seen to decrease continuously with time.

The quantity $n_e$ may be determined experimentally as a function of time throughout the night (e.g. MITRA, 1957a). The quantity $\alpha_a$ can be determined immediately from these measurements, and $\alpha_b$ may then be found from (8).

If only two species of ion are present, their relative concentrations may be eliminated from (5) and (9) to give

$$\alpha_b^2 = (\alpha_1 + \alpha_2)\alpha_a - \alpha_1\alpha_2. \tag{10}$$

A graph of $\alpha_b^2$ against $\alpha_a$, for different times on the same night, will therefore give the values of $\alpha_1$ and $\alpha_2$, the recombination coefficients of the two species of ion. Since, however, the double differential of the electron density curve is involved, the accuracy of determination of $\alpha_1$ and $\alpha_2$ may be low (particularly if $\alpha_1$ and $\alpha_2$ are of the same order of magnitude).

A more useful analysis, for the case of two ionic species, can be made using a new variable $G$,

$$G = \alpha_b^2/\alpha_a^2. \tag{11}$$

This quantity, which can be readily determined experimentally, is never less than unity, if the recombination process accords with the assumptions made in this section. Its value is given by

$$G = \frac{n_1\alpha_1^2 + n_2\alpha_2^2}{(n_1\alpha_1 + n_2\alpha_2)^2}. \tag{12}$$

Suppose that initially the ionized region consists predominantly of the more rapidly combining ion, so that $\alpha_1 > \alpha_2$, $n_1\alpha_1 \gg n_2\alpha_2$. Then $G$ will be nearly unity. As the recombination proceeds, the proportion of ion 2 increases, until finally $n_2\alpha_2^2 \gg n_1\alpha_1^2$, when $G$ is once again unity. At intermediate times, $G$ passes through a maximum value,

$$G_m = \frac{(\alpha_1 + \alpha_2)^2}{4\alpha_1\alpha_2} \text{ when } \frac{n_1}{n_2} = \frac{\alpha_2}{\alpha_1}. \tag{13}$$

As an example, the data used by MITRA (1957a, Fig. 2) will be considered. He found values of $\alpha_a$ (for the $E$-layer at Watheroo) which decreased from about $10^{-8}$ to less than $10^{-9}$ cm$^3$ sec$^{-1}$ as the night progressed. Since the determination of $\alpha_a$ is, as he points out, subject to considerable error due to vertical motions when $\alpha$ becomes small, his data for $\alpha_a$ are taken only from 0·5 to 4·5 hr after sunset. Over this period, he fits a straight line with

$$\frac{d\alpha_a}{dt} = 4{\cdot}72 \times 10^{-13} \text{ cm}^3 \text{ sec}^{-2}$$

and

$$\alpha_a = 6{\cdot}8 \times 10^{-9} \text{ cm}^3 \text{ sec}^{-1}, \quad n_e = 9 \times 10^3 \text{ cm}^{-3}$$

at 2·5 hr after sunset, the centre of the period considered. Using (8) and (11),

$$\alpha_b = 0{\cdot}99 \times 10^{-8} \text{ cm}^3 \text{ sec}^{-1}$$
$$G = 2{\cdot}14$$

From (13), $G_m = 2{\cdot}14$ for $\alpha_1/\alpha_2 = 6{\cdot}4$. It follows that the two recombination coefficients required to explain the observed behaviour must differ by at least this factor.

The set of equations (2) may be solved, at least in principle, by dividing each equation by one of the set:

$$\frac{dn_r}{dn_1} = \frac{\alpha_r}{\alpha_1}\frac{n_r}{n_1}.$$

Integrating, and allowing the additional suffix $i$ to stand for the initial condition,

$$\begin{aligned} n_r &= n_{ri}(n_1/n_{1i})^{\alpha_r/\alpha_1} \\ n_e &= \sum_{r=1}^{m} n_{ri}(n_1/n_{1i})^{\alpha_r/\alpha_1} \end{aligned} \tag{14}$$

The differential equation for $n_1$ may be written

$$\frac{dn_1}{dt} = \alpha_1 n_1 \sum_{r=1}^{m} n_{ri}(n_1/n_{1i})^{\alpha_r/\alpha_1} \tag{15}$$

and the variables are now separable. Analytic quadrature of this equation cannot, however, be performed in the general case.



VOL. 20 1961

To summarize, it has been shown that the variation of effective nocturnal recombination coefficient $\alpha_a$ can be satisfactorily explained on the hypothesis that two species of ion are present, with recombination coefficients differing by a factor of at least six.

## 3. Eclipse Recombination

In principle, study of the ionization density variation during a solar eclipse should be an ideal method for the determination of ionospheric recombination coefficients. Measurements on control days preceding and following the eclipse day give the normal variation of the electron density, and the electron production on the eclipse day may be deduced by multiplying the production on the control days by a factor corresponding to the geometrical obscuration of the sun's disk, at the height of the ionized layer in question. Application of (1) then gives the value of $\alpha$ throughout the eclipse.

Ratcliffe (1956) has indicated serious difficulties associated with this simple approach. Typically, between first and second contact, (1) gives a value of $\alpha$ that may be as high as $3 \times 10^{-8}$ cm$^3$ sec$^{-1}$ for the $E$-layer; between third and fourth contact, a lower value of $\alpha$ is nearly always observed, sometimes as low as $0{\cdot}5 \times 10^{-8}$ cm$^3$ sec$^{-1}$. In addition, the minimum value of electron density is nearly always higher than would be expected on a simple theory.

Attempts have been made (Szendrei and McElhinny, 1956; Minnis, 1956, 1958a; McElhinny, 1959) to explain the failure of (1) by assuming a non-uniform distribution of the ionizing radiation over the sun's disk. In every case, however, where this analysis has been satisfactorily performed, with a constant value of $\alpha$, it has led to a model of the solar brightness distribution which has a bright area on the western limb. Since there is no reason to suppose that the western limb of the sun is consistently brighter than the eastern, it would seem that some variation in the value of $\alpha$ throughout the eclipse is inevitable. A detailed discussion is given by McElhinny (1959).

If, once again, an ionosphere consisting of $m$ species of positive ion is considered, it is evident that the value of $\alpha$ determined from (1) is in fact $\alpha_a$, as defined in Section 2. The value of $\alpha_a$ depends on the relative abundance of the various species; so a value deduced for the initial phase of the eclipse corresponds to the relative abundance of these ions under control-day conditions. As the eclipse progresses, just as at night, the value of $\alpha_a$ decreases, reaching its minimum value in the final phase of the eclipse. The hypothesis of several ionic species therefore explains at least the gross features of the variation of $\alpha$.

McElhinny (1959) has fitted measurements of $E$-layer ionization in South Africa in the solar eclipse of 25 December 1954, using a two-species ion model. He chose two models; in the notation used in this paper, the initial conditions, at first contact in the eclipse, correspond to

$$n_1 = 0{\cdot}8n_e, \quad n_2 = 0{\cdot}2n_e$$

$$\alpha_1 = 5 \times 10^{-9} \text{ cm}^3 \text{ sec}^{-1}, \quad \alpha_2 = 10^{-7} \text{ cm}^3 \text{ sec}^{-1}$$

giving

$$\alpha_a = 2{\cdot}4 \times 10^{-8} \text{ cm}^3 \text{ sec}^{-1}, \quad \alpha_b = 4{\cdot}7 \times 10^{-8} \text{ cm}^3 \text{ sec}^{-1}$$

while the second model corresponds to

$$n_1 = n_2 = 0{\cdot}5n_e$$

$$\alpha_1 = 10^{-9}\ \text{cm}^3\ \text{sec}^{-1}, \quad \alpha_2 = 10^{-7}\ \text{cm}^3\ \text{sec}^{-1}$$

giving

$$\alpha_a = 5 \times 10^{-8}\ \text{cm}^3\ \text{sec}^{-1}, \quad \alpha_b = 7{\cdot}1 \times 10^{-8}\ \text{cm}^3\ \text{sec}^{-1}.$$

The large magnitude of $\alpha_a$ for the initial phase of the eclipse will be noted. One would expect that the value of $\alpha_a$ would be the same for the initial phase of the eclipse, whatever particular model of ionic recombination coefficients were chosen. In fact, examination of his Fig. 3 indicates that the best fit would be obtained from a value of $\alpha_a$ intermediate between the two; near to $3 \times 10^{-8}\ \text{cm}^3\ \text{sec}^{-1}$.

Values of the quantity $G$, which may be taken to represent the degree of diversity between the ions composing the ionized layer, may be readily calculated from his data; his two models quoted above give $G = 3{\cdot}8$ and $2{\cdot}0$ respectively, for the initial phase of the eclipse.

Further light may be shed on this problem by considering the minimum electron density occurring during the eclipse. If two points on the electron-density–time curve can be identified at which the gradient is zero: one near the commencement of the eclipse, and the other at the minimum electron density, one may write the relations

$$q_{ei} = n_{ei} \sum_{r=1}^{m} \alpha_r n_r = {n_{ei}}^2 \alpha_{ai} \tag{16}$$

$$q_{em} = {n_{em}}^2 \alpha_{am} \tag{17}$$

where the suffixes $i$ and $m$ denote the initial and minimum electron density points. Dividing the two equations, and defining a new ratio $R$,

$$R = \frac{q_{em}}{q_{ei}} \cdot \frac{{n_{ei}}^2}{{n_{em}}^2} = \frac{\alpha_{am}}{\alpha_{ai}}. \tag{18}$$

As pointed out by RATCLIFFE (1956), one would expect $R$ to be unity on a simple model, with a single recombination process. If several species of ion are present, however, $R$ must be less than unity under all conditions. Inspection of the results quoted by Ratcliffe and subsequent workers, indicate that in every case $R$ is less than 1; values ranging from 0·35 to 0·8 have been recorded.

This has hitherto been interpreted as meaning that the electron production between second and third contact is much greater than would be supposed from geometrical considerations; for instance, a substantial amount of ionizing coronal emission might be present. However, the departure of $R$ from unity may be completely accounted for by (18). This is illustrated by the results of MCELHINNY (1959); these give $R = 0{\cdot}61$, corresponding to a 39 per cent decrease in effective recombination coefficient during the time between first and second contact.

An analogous equation to (8) will now be derived, for the case where $q \neq 0$. As in (7),

$$\frac{d(\alpha_r n_r/n_e)}{dt} = \frac{\alpha_r}{n_e}\frac{dn_r}{dt} - \frac{\alpha_r n_r}{{n_e}^2}\frac{dn_e}{dt}. \tag{19}$$

Writing

$$\frac{dn_r}{dt} = q_r - \alpha_r n_r n_e \tag{20}$$

where $q_r$ is the rate of production of the $r$th ion species, and using the relation

$$\frac{dn_e}{dt} = q_e - \alpha_a {n_e}^2 \tag{21}$$

where $q_e = \sum_{n=1}^{m} q_r$, and, substituting (20) and (21) into (19),

$$\frac{d(\alpha_r n_r/n_e)}{dt} = \frac{\alpha_r q_r}{n_e} - \alpha_r {}^2 n_r + \left(\alpha_a - \frac{q_e}{{n_e}^2}\right)\alpha_r n_r. \tag{22}$$

Before summing this relation over $r$, the $q_r$ must be determined. It is assumed that the $q_r$ are always in a fixed ratio to $q_e$ throughout the eclipse; so if the suffix $j$ denotes a situation of complete photo-equilibrium prior to the eclipse, from (20) and (21),

$$\left.\begin{aligned} q_{rj} &= \alpha_r n_{rj} n_{ej} \\ q_{ej} &= \alpha_{aj} {n_{ej}}^2 \\ \frac{q_r}{q_e} &= \frac{q_{rj}}{q_{ej}} \end{aligned}\right\} \tag{23}$$

Substituting (23) into (22) and summing over $r$,

$$\frac{d\alpha_a}{dt} = \frac{q_e}{q_{ej}}\frac{n_{ej}}{n_e}\sum_{r=1}^{m} \alpha_r {}^2 n_{rj} - \sum_{r=1}^{m} \alpha_r {}^2 n_r + \left(\alpha_a - \frac{q_e}{q_{ej}}\frac{{n_{ej}}^2}{{n_e}^2}\alpha_{aj}\right)\sum_{r=1}^{m} \alpha_r n_r.$$

With the definitions of $\alpha_a$ and $\alpha_b$, this becomes

$$\frac{d\alpha_a}{dt} = \frac{q_e}{q_{ej}}\frac{{n_{ej}}^2}{n_e}\alpha^2{}_{bj} - n_e {\alpha_b}^2 + n_e {\alpha_a}^2 - \frac{q_e}{q_{ej}}\frac{{n_{ej}}^2}{n_e}\alpha_{aj}\alpha_a$$

$$\frac{d\alpha_a}{dt} = n_e({\alpha_a}^2 - {\alpha_b}^2) + \frac{q_e}{q_{ej}}\frac{{n_{ej}}^2}{n_e}({\alpha_{bj}}^2 - \alpha_{aj}\alpha_a) \tag{24}$$

which is the required extension of (8) to eclipse recombination. It should be noted that this equation, like (8), does not depend on any assumption about the number of ionic species involved.

## 4. Diurnal Recombination

Appleton (1937, 1955) has described a method, particularly applicable to the $E$-layer of the ionosphere, for determining the recombination coefficient of the ionization. It depends on what he has called the "sluggishness" of the ionosphere; the fact that the electron density in a layer lags behind the rate of electron production. In particular, the maximum electron density at a given height does not occur simultaneously with the maximum electron production rate (which



VOL. 20 1961

occurs at local noon). He has shown that, under certain assumptions, the time difference between these two maxima is

$$\tau = \frac{1}{2\alpha_s n_e}. \tag{25}$$

Values of $\alpha_s$ determined from this equation usually lie around $10^{-8}$ cm$^3$ sec$^{-1}$; much less than the value of $\alpha_a$ found during the initial phase of an eclipse.

It is shown here that, in the presence of an ionosphere consisting of a mixture of ions, the value of $\alpha_s$ deduced from (25) would be expected to be different, and less than, the value of $\alpha_a$ deduced from eclipse measurements.

In order to carry out Appleton's analysis for several ionic species, the starting point is taken as equation (20):

$$\frac{dn_e}{dt} = q_r - \alpha_r n_r n_e. \tag{20}$$

Assuming, as in Section 3, that the rates of production $q_r$ of the ionic species vary proportionally, they may be written

$$q_r = K_r q_e \tag{26}$$

and (20) becomes

$$\frac{dn_r}{dt} = K_r q_e - \alpha_r n_r n_e. \tag{27}$$

Following Appleton, it is assumed, as a first approximation, that the left-hand side of (27) is negligible. Then, approximately,

$$n_r = K_r q_e / \alpha_r n_e. \tag{28}$$

Summing (28) over all $r$,

$$n_e^2 = q_e \sum_{r=1}^{m} K_r/\alpha_r. \tag{29}$$

Defining $\alpha_c$ by

$$\alpha_c^{-1} = \sum_{r=1}^{m} K_r/\alpha_r \tag{30}$$

the approximate value of $n_e$ given by (29) may be substituted in (28);

$$n_r = \frac{K_r q_e^{1/2}}{\alpha_c \alpha_r}. \tag{31}$$

Substitution of (31) in the left-hand side of (27) enables a more accurate determination of $n_r$ to be made:

$$n_r = \frac{K_r q_e}{\alpha_r n_e} - \frac{K_r \alpha_c^{1/2} \dot{q}_e}{2\alpha_r^2 n_e q_e^{1/2}} \tag{28a}$$

and on summing this equation

$$n_e^2 = \frac{q_e}{\alpha_c} - \frac{\alpha_c^{1/2} \dot{q}_e}{2 q_e^{1/2}} \sum_{r=1}^{m} K_r \alpha_r^{-2}. \tag{29a}$$



VOL. 20 1961

Finally, defining $\alpha_d$ by

$$\alpha_d^{-2} = \sum_{r=1}^{m} K_r/\alpha_r^2 \tag{30a}$$

the second approximation to the electron density is

$$n_e^2 = \frac{q_e}{\alpha_c} - \frac{\alpha_c^{1/2}}{2\alpha_d^2}\frac{\dot{q}_e}{q_e}. \tag{32}$$

Writing the electron production $q_e$ in the form

$$q_e = q_{e0}(1 - ft^2) \tag{33}$$

where $t$ is time measured from local noon in seconds, and $f$ is a constant (a form which applies for times near noon), (32) becomes

$$n_e^2 = \frac{q_{e0}}{\alpha_c} - \frac{q_{e0}ft^2}{\alpha_c} + \frac{q_{e0}ft\alpha_c^{1/2}}{\alpha_d^2}. \tag{34}$$

Differentiation of (34) shows that the maximum value of $n_e$ occurs when

$$\tau = \frac{\alpha_c}{2\alpha_d^2 n_e}. \tag{35}$$

Thus it is seen that the value of $\alpha_s$ measured by (25) is just

$$\alpha_s = \alpha_d^2/\alpha_c \tag{36}$$

which would be equal to the $\alpha_a$ of the previous sections if only one ion species were present.

Using, as in Section 3, the suffix $j$ to denote photochemical equilibrium, (30) and (30a) can be rewritten as

$$\alpha_c^{-1} = \sum_{r=1}^{m} n_{ej}n_{rj}/q_e = n_{ej}^2/q_e = \alpha_{aj}^{-1} \tag{37}$$

$$\alpha_d^{-1} = \sum_{r=1}^{m} n_{aj}n_{rj}/q_e\alpha_r = (n_{ej}\alpha_{aj})^{-1}\sum_{r=1}^{m} n_{rj}/\alpha_r \tag{38}$$

and consequently

$$\alpha_s^{-1} = \sum_{r=1}^{m} (n_{rj}/n_{ej})\alpha_r^{-1}. \tag{39}$$

Comparing (39) with (5), it is evident that the recombination coefficient measured by the "sluggishness" of the ionosphere is just the harmonic mean for all ions present, weighted by their relative abundance, under conditions of photochemical equilibrium.

As is well known, the harmonic mean is always less than the arithmetic mean; so $\alpha_s \leqslant \alpha_a$ under all conditions. Experimental values of $\alpha_s$ are usually found (APPLETON, 1953) to be about $10^{-8}$ cm$^3$ sec$^{-1}$, for the $E$-layer. Assuming that the difference between this value for $\alpha_s$ and the value for $\alpha_a$ of $3 \times 10^{-8}$ cm$^3$ sec$^{-1}$ found in Section 3 is due to the presence of just two species of ion, it is

possible to calculate the relationship between the two recombination coefficients, $\alpha_1$ and $\alpha_2$. Writing (5) and (39) in the form

$$\alpha_a = \frac{n_1}{n_e}\alpha_1 + \frac{n_2}{n_e}\alpha_2 \tag{40}$$

$$\alpha_s^{-1} = \frac{n_1}{n_e}\alpha_1^{-1} + \frac{n_2}{n_e}\alpha_2^{-1} \tag{41}$$

and eliminating $n_1$ and $n_2$, one arrives at

$$\frac{\alpha_2}{\alpha_1} = \frac{\alpha_a - \alpha_1}{\alpha_s - \alpha_1} \cdot \frac{\alpha_s}{\alpha_1} \tag{42}$$

assuming that $\alpha_1$ is the smaller recombination coefficient. If $\alpha_a$ and $\alpha_s$ are regarded as fixed, this recombination coefficient ratio has a minimum value of

$$\left(\frac{\alpha_2}{\alpha_1}\right)_{\min} = 2\sqrt{\left[\left(\frac{\alpha_a}{\alpha_s}\right)^2 - \left(\frac{\alpha_a}{\alpha_s}\right)\right]} + \frac{2\alpha_a}{\alpha_s} - 1. \tag{43}$$

Equation (43) therefore places a minimum value on the recombination coefficient ratio for any experimental values of $\alpha_a$ and $\alpha_s$. Inserting the values stated above,

$$\left(\frac{\alpha_2}{\alpha_1}\right)_{\min} = 9{\cdot}92.$$

Confirming the results of the previous sections, therefore, two widely different recombination coefficients are required to explain the observed results.

As to the precise values of $\alpha_1$ and $\alpha_2$ involved, one may assume that the values used in (40) to (43) are the same as those dealt with in Section 2, relating to nocturnal recombination. If this is the case, by an argument similar to the above, one may use the values of $\alpha_a$ and $\alpha_b$ to obtain the following equation for the night:

$$\alpha_2 = \frac{0{\cdot}98 \times 10^{-8} - 0{\cdot}68\alpha_1}{0{\cdot}68 \times 10^{-8} - \alpha_1} \times 10^{-8} \tag{44}$$

while for the day

$$\alpha_2 = \frac{3 \times 10^{-8} - \alpha_1}{10^{-8} - \alpha_1} \times 10^{-8}. \tag{45}$$

Combining (44) and (45) gives the values

$$\alpha_1 = 6 \times 10^{-9}\ \mathrm{cm^3\ sec^{-1}}, \quad \alpha_2 = 6 \times 10^{-8}\ \mathrm{cm^3\ sec^{-1}}.$$

Corresponding to these values are relative concentrations of species 2, the rapidly recombining ion, of 0·47 in the day and 0·02 at night.

## 5. Discussion

In the previous sections it has been shown that some of the discrepant results obtained from measurements of effective recombination coefficient can be explained by assuming that several types of ion are present, each recombining



VOL. 20 1961

separately with its own recombination coefficient. Possibilities for the production of these types of ion will now be briefly examined.

BATES and NICOLET (1960) have suggested that the following are the important recombination processes in the ionosphere:

$$NO^+ + e \rightarrow N' + O'$$
$$O_2^+ + e \rightarrow O' + O''$$

which they assume to have rate coefficients, respectively, of about $3 \times 10^{-9}$ and $3 \times 10^{-8}$ cm$^3$ sec$^{-1}$. The nitric oxide ions are assumed to be produced by ion–atom interchange between molecular nitrogen and atomic oxygen ions:

$$O^+ + N_2 \rightarrow NO^+ + N + 1{\cdot}1 \text{ eV}.$$

Ions of molecular oxygen and nitric oxide have in fact been detected at *E*-layer heights by JOHNSON *et al.* (1958). They found virtually no molecular nitrogen ions; this may be due to the very rapid dissociative recombination of molecular nitrogen ions.

One may therefore provisionally identify the two ion species treated in the previous section as nitric oxide and molecular oxygen. Relative abundances of these two ions found by JOHNSON *et al.* at a height of 110 km in the arctic ionosphere are given in Table 1, together with those deduced in Section 4.

Table 1.

| | $NO^+$ | $O_2^+$ |
|---|---|---|
| Rocket (night) | 0·97 | 0·03 |
| Section 4 | 0·98 | 0·02 |
| Rocket (day) | 0·73 | 0·27 |
| Section 4 | 0·53 | 0·47 |

In view of the assumptions involved, the agreement must be regarded as satisfactory.

With the deduced values of recombination coefficient, the production rates for the two species may be deduced from (23). For an electron concentration of $10^5$ cm$^{-3}$, the respective rates for nitric oxide and molecular oxygen ions are 31 and 270 cm$^3$ sec$^{-1}$.

One further point relates to the possibility of part of the nocturnal recombination process being due to the formation of negative ions. MITRA (1957b) has studied this problem and concludes that, at 110 km height, the ratio of negative to positive ions never exceeds 4 per cent. On this basis, it is felt that the contribution of negative ions to the recombination can be neglected.

## 6. CONCLUSIONS

It has been demonstrated that three methods for determining effective recombination coefficient in the ionosphere, all in current use, may have very different values if the ionosphere consists of more than one species of ion. Results

obtained by comparing effective recombination coefficients derived by the three methods are consistent with rocket measurements of ionospheric composition.

It is clear that many more detailed studies must be carried out using the equations derived in this paper before the recombination coefficients of the ionic species can be said to have been accurately determined.

## References


| | | |
|---|---|---|
| Appleton E. V. | 1937 | *Proc. Roy. Soc.* A **162,** 451. |
| Appleton E. V. | 1953 | *J. Atmosph. Terr. Phys.* **3,** 282. |
| Appleton E. V. and Lyon A. J. | 1955 | *Physics of the Ionosphere,* p. 20. Physical Society, London. |
| Appleton E. V. and Lyon A. J. | 1957 | *J. Atmosph. Terr. Phys.* **10,** 1. |
| Bates D. R. and McDowell M. R. C. | 1957 | *J. Atmosph. Terr. Phys.* **10,** 96. |
| Bates D. R. and Nicolet M. | 1960 | *J. Atmosph. Terr. Phys.* **18,** 65. |
| Johnson C. Y., Meadows E. B. and Holmes J. C. | 1958 | *J. Geophys. Res.* **63,** 443. |
| McElhinny M. W. | 1959 | *J. Atmosph. Terr. Phys.* **14,** 273. |
| Minnis C. M. | 1956 | *Solar Eclipses and the Ionosphere,* p. 158. Pergamon Press, London. |
| Minnis C. M. | 1958a | *J. Atmosph. Terr. Phys.* **12,** 266. |
| Minnis C. M. | 1958b | *J. Atmosph. Terr. Phys.* **12,** 272. |
| Mitra A. P. | 1957a | *J. Atmosph. Terr. Phys.* **10,** 140. |
| Mitra A. P. | 1957b | *J. Atmosph. Terr. Phys.* **10,** 153. |
| Ratcliffe J. A. | 1956 | *Solar Eclipses and the Ionosphere,* p. 1. Pergamon Press, London. |
| Ratcliffe J. A. | 1960 | *Physics of the Upper Atmosphere,* p. 378. Academic Press, New York. |
| Szendrei M. E. and McElhinny M. W. | 1956 | *Solar Eclipses and the Ionosphere,* p. 74. Pergamon Press, London. |




VOL. 20 1961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 31 to 39. Pergamon Press Ltd. Printed in Northern Ireland

# Analytical considerations of ionospheric windows for low frequency radio waves

J. HEADING
The University of Southampton

(*Received* 17 *August* 1960)

**Abstract**—After a brief consideration of an isotropic ionosphere, the reflection coefficients for anisotropic ionospheres with free space below and above and examined for their zeros. The particular case studied corresponds to vertical propagation in an Epstein distribution of ionization density together with a vertical magnetic field. In this case it is found that zeros are possible only for one particular polarized wave in the absence of a collision frequency, and the results are illustrated with reference to numerical data from a physical ionosphere.

## 1. INTRODUCTION

SEVERAL years ago, HINES (1953) published a new method for solving problems involving wave propagation in a varying medium for which the reflection coefficients were required. One of the interesting results of his paper was to produce the profiles of several media for which the reflection coefficient vanished at a discrete set of frequencies yielding perfect transmission through the regions at those frequencies, thereby forming a series of windows in the medium with respect to frequency. Moreover, it has recently been stressed in the NASA SPACE SCIENCE PROGRAM (1960) that the probable existence of very low frequency windows in the ionosphere would prove useful for interplanetary communication. In this present paper, we intend to consider this subject analytically.

The differential equation with which HINES was concerned was not exactly the type applicable to an isotropic ionosphere. His refractive index was independent of the frequency, while it is well known that the refractive index for the ionosphere is a function of the frequency. We commence our study by considering the case of vertical incidence on an isotropic ionosphere in which the collision frequency is zero. The appropriate equation for the horizontal component $E$ of the electric field vector is given by

$$(d^2E/dz^2) + k^2(1 - X)E = 0 \tag{1}$$

where $k = \omega/c$, the assumed time factor being $e^{i\omega t}$. Here, $X$ is proportional to the electron density at the height $z$, but dependent upon the frequency in that it contains a factor $k^2$ in its denominator. As far as the physical ionosphere is concerned, such an equation is quite inappropriate in the case of long and very long waves, but since its solution forms the basis of calculations for an anisotropic ionosphere in which both the direction of propagation and the direction of the externally applied magnetic field are vertical, the simpler equation may first of all be considered.

## 2. TRANSMISSION THROUGH AN ISOTROPIC HOMOGENEOUS SLAB

If the wave is incident normally upon a uniform isotropic slab of thickness $h$

lying between the planes $z = 0$ and $z = h$, such that $X = k_0^2/k^2$ where $k_0$ is constant, the reflection and transmission coefficients may easily be calculated. If $n = \sqrt{(1 - X)}$, we obtain

$$R = \frac{(n^2 - 1)(1 - e^{2iknh})}{(n + 1)^2 e^{2iknh} - (n - 1)^2} \tag{2}$$

$$T = \frac{4ne^{iknh}e^{ikh}}{(n + 1)^2 e^{2iknh} - (n - 1)^2} \tag{3}$$

where $|R|^2 + |T|^2 = 1$. The branch chosen for $n$ in both underdense ($X < 1$) and overdense ($X > 1$) models is irrelevant, since the values of $R$ and $T$ are independent of the sign of $n$.

The reflection coefficient vanishes only when $e^{2iknh} = 1$, or when

$$2iknh = 2\pi iN$$

where $N$ is any integer, provided $n$ is real. Hence, upon squaring,

$$k^2(1 - X)h^2 = \pi^2 N^2$$

or

$$(k^2 - k_0^2)h^2 = \pi^2 N^2$$

with the solution

$$k = \sqrt{\left(k_0^2 + \frac{\pi^2 N^2}{h^2}\right)}.$$

Now the frequency $f$ is given by $f = ck/2\pi$, so in terms of the critical frequency $f_0$, we obtain

$$f = \sqrt{\left(f_0^2 + \frac{c^2 N^2}{4h^2}\right)}. \tag{4}$$

The corresponding value of $T$ equals $(-1)^N e^{ikh}$, of unit modulus. This discrete set of values of $k$ ensures that $n$ is real, but for all integral values of $N$ we note that $k > k_0$. We shall not describe this model as yielding a low frequency window; this term will be reserved for frequencies less than the critical frequency. For the homogeneous isotropic slab, then, no low frequency windows are possible.

## 3. Transmission through an Isotropic Epstein Ionosphere

If we place

$$X = c_1 + \frac{[(c_2 - c_1)(e^{\alpha z} - 1) + 4a_3]e^{\alpha z}}{(e^{\alpha z} + 1)^2}$$

choosing $c_1$, $c_2$ and $a_3$ so that $X \geqslant 0$ for all $z$, we obtain what may be referred to as the Epstein (1930) profile for the ionosphere. It is understood that the constants $c_1$, $c_2$ and $a_3$ contain a factor $k^2$ in their denominators, and that $\alpha$ is independent of $k$. Above the ionosphere $X \to c_2$ while below $X \to c_1$. If we write

$$1 - c_1 = a_1^2, \qquad 1 - c_2 = a_2^2$$

$$M = \sqrt{\left(\frac{1}{4} - \frac{4k^2 a_3}{\alpha^2}\right)} \qquad \text{when} \qquad \frac{1}{4} \geqslant \frac{4k^2 a_3}{\alpha^2}$$



VOL
20
196

$$P = \sqrt{\left(\frac{4k^2a_3}{\alpha^2} - \frac{1}{4}\right)} \qquad \text{when} \qquad \frac{1}{4} \leqslant \frac{4k^2a_3}{\alpha^2}$$

the work of EPSTEIN shows that the moduli of the reflection and transmission coefficients for incidence below the ionosphere are given by the following results:

$$|R|^2 = \frac{\cosh\left[2\pi k(a_1 - a_2)/\alpha\right] + \cos 2\pi M}{\cosh\left[2\pi k(a_1 + a_2)/\alpha\right] + \cos 2\pi M} \tag{5}$$

or

$$\frac{\cosh\left[2\pi k(a_1 - a_2)/\alpha\right] + \cosh 2\pi P}{\cosh\left[2\pi k(a_1 + a_2)/\alpha\right] + \cosh 2\pi P} \tag{6}$$

$$|T|^2 = \frac{a_1}{a^2} \cdot \frac{\cosh\left[2\pi k(a_1 + a_2)/\alpha\right] - \cosh\left[2\pi k(a_1 - a_2)/\alpha\right]}{\cosh\left[2\pi k(a_1 + a_2)/\alpha\right] + \cos 2\pi M} \tag{7}$$

or

$$\frac{a_1}{a^2} \cdot \frac{\cosh\left[2\pi k(a_1 + a_2)/\alpha\right] - \cosh\left[2\pi k(a_1 - a_2)/\alpha\right]}{\cosh\left[2\pi k(a_1 + a_2)/\alpha\right] + \cosh 2\pi P}. \tag{8}$$

Since the hyperbolic cosine is always greater or equal to unity when its argument is real it is obvious that $|R|$ can never vanish unless $a_1 = a_2$. Moreover, in order to have free space below the ionosphere, we must have $c_1 = 0$ so $a_1 = 1$. The profile for $X$ then becomes

$$X = \frac{4a_3e^{\alpha z}}{(e^{\alpha z} + 1)^2} = a_3 \operatorname{sech}^2 \frac{\alpha z}{2} \tag{9}$$

a symmetrical profile attaining its maximum value $a_3$ when $z = 0$. Formula (5) alone is applicable, yielding

$$|R|^2 = \frac{1 + \cos 2\pi M}{\cosh\left(4\pi k/\alpha\right) + \cos 2\pi M} \tag{10}$$

with $1 = |R|^2 + |T|^2$. For this particular model, $|R|$ vanishes only if $\cos 2\pi M = -1$, namely when

$$2\pi M = j\pi$$

where $j$ is any odd integer. Then

$$\frac{j^2}{4} = M^2 = \tfrac{1}{4} - \frac{4k^2a_3}{\alpha^2}$$

or

$$\frac{4k^2a_3}{\alpha^2} = \frac{(1 - j^2)}{4}. \tag{11}$$

Since $a_3$ contains implicitly a factor $k^2$ in its denominator, this equation is really independent of $k$; moreover for the ionosphere $a_3$ must be positive, so no value of the odd integer $j$ exists satisfying (11). We conclude that $|R|$ can never vanish for an isotropic Epstein ionosphere.

If on the other hand we abandon the ionospheric interpretation, and regard the function $1 - X$ merely as the square of a frequency-independent refractive index for which $a_3$ can take negative values, we find, placing $a_3 = -b_3$ say where $b_3$ is independent of $k$, that the model governed by the equation



VOL. 20 1961

$$\frac{d^2E}{dz^2} + k^2\left[1 + \frac{4b_3e^{\alpha z}}{(e^{\alpha z} + 1)^2}\right]E = 0$$

possesses zero reflection coefficient whenever

$$k^2 = \frac{(j^2 - 1)\alpha^2}{16b_3}$$

$j$ being an odd integer. A discrete set of frequencies is thereby produced yielding windows through the medium at these values.

## 4. Anisotropic Models

If a vertical magnetic field is superimposed on the above models, the equations for the two components $E_x$ and $E_y$ of the electric field perpendicular to the direction of propagation have been given, for example, by CLEMMOW and HEADING (1954); they take the forms

$$\left.\begin{aligned}\frac{d^2F}{dz^2} + k^2\left(1 - \frac{X}{1 - Y}\right)F = 0\\ \frac{d^2G}{dz^2} + k^2\left(1 - \frac{X}{1 + Y}\right)G = 0\end{aligned}\right\} \tag{12}$$

where $F = E_x + iE_y$ and $G = E_x - iE_y$. Here, $Y$ is proportional to the strength of the external magnetic field, and contains a factor $k$ in its denominator. The combinations $F$ and $G$ are propagated independently, representing circularly-polarized waves.

When free space exists below the ionosphere, let the solution for $F$ there be $e^{-ikz} + R_Fe^{ikz}$ and that for $G$ be $e^{-ikz} + R_Ge^{ikz}$, where $R_F$ and $R_G$ are respectively the reflection coefficients for the two circularly polarized waves. Taking arbitrary combinations of these two fields, we obtain

$$E_x = \frac{(A + B)e^{-ikz}}{2} + \frac{(AR_F + BR_G)e^{ikz}}{2}$$

$$iE_y = \frac{(A - B)e^{-ikz}}{2} + \frac{(AR_F - BR_G)e^{ikz}}{2}.$$

If no incident $E_y$ field is present, we place $A = B$, yielding for the reflection and conversion coefficients associated with an incident field polarized in the $x$-direction

$$R_{xx} = \frac{(R_F + R_G)}{2}, \quad R_{xy} = \frac{i(R_G - R_F)}{2}. \tag{13}$$

Similarly, if no incident $E_x$ field is present, we place $A = -B$, yielding

$$R_{yy} = \frac{(R_F + R_G)}{2}, \quad R_{yx} = \frac{i(R_F - R_G)}{2}. \tag{14}$$

Hence $R_{xx}$ and $R_{yy}$ can only vanish provided $R_F = -R_G$, and the conversion coefficients $R_{xy}$ and $R_{yx}$ can only vanish provided $R_F = R_G$.



VOL. 20 1961

The values of $R_F$ and $R_G$ may be obtained from the reflection coefficients already quoted for the isotropic models. For $R_F$, we use the value of $R$ but with the constant factor appearing in $X$ modified by the divisor $1 - Y$; for $R_G$ we modify $R$ by using the divisor $1 + Y$. This procedure is justified merely by inspecting the equations (12).

## 5. The Anisotropic Homogeneous Slab

The reflection coefficient $R_F$ is given by (12), with

$$n = \sqrt{\left(1 - \frac{X}{1 - Y}\right)} \tag{15}$$

while $R_G$ is produced when

$$n = \sqrt{\left(1 - \frac{X}{1 + Y}\right)}. \tag{16}$$

The reflection coefficient vanishes when $n$ is real and when

$$knh = \pi N$$

where $N$ is a positive integer. If we place $X = k_0^2/k^2$, $Y = k_1/k$ with $k_1 > 0$, we obtain when $R_F$ vanishes

$$k^2\left[1 - \frac{k_0^2/k^2}{1 - (k_1/k)}\right] = \frac{\pi^2 N^2}{h^2}$$

or

$$f(k) \equiv k^3 - k_1 k^2 - \left(k_0^2 + \frac{\pi^2 N^2}{h^2}\right)k + \frac{\pi^2 N^2 k_1}{h^2} = 0$$

a cubic equation in $k$.

If the graph of $f(k)$ against $k$ is plotted, it will be noticed that $f(0) > 0, f(+\infty) > 0$, $f(-\infty) < 0$, $f(k_1) < 0$. Hence two positive roots must exist for all $N$, one in the range $0 < k < k_1$ and the other in the range $k > k_1$. Moreover, when $N = 0$, $k = 0$ and as $N \to \infty$, $k \to k_1$ as far as the range $0 < k < k_1$ is concerned. A discrete set of values of $k$ has therefore been found below the gyro-frequency. If $N$ is small enough, $f(k_0) < 0$, showing that values of $k < k_0$ exist however overdense the medium may be in the isotropic sense. If $k_1 > k_0$, $f(k_0)$ is positive for larger values of $N$, showing that the majority of the values of $k$ must lie in the range $k_0 < k < k_1$.

The reflection coefficient $R_G$ vanishes when

$$k^2\left[1 - \frac{k_0^2/k^2}{1 + (k_1/k)}\right] = \frac{\pi^2 N^2}{h^2}$$

namely when

$$g(k) \equiv k^3 + k_1 k^2 - \left(k_0^2 + \frac{\pi^2 N^2}{h^2}\right)k - \frac{\pi^2 N^2 k_1}{h^2} = 0.$$

Since $g(0) < 0$, $dg(0)/dk < 0$, $g(+\infty) > 0$, there can only be one positive value of $k$ satisfying this equation for any given $N$. When $N = 0$, we note that $g(k_0) > 0$, showing that there is at least one value of $k$ below its isotropic critical value, but

when $N$ is large, all values of $k$ will be much greater than $k_0$, since $g(k_0)$ is then large and negative.

## 6. The Anisotropic Epstein Ionosphere

Using (11), we deduce that the reflection coefficient $R_F$ vanishes when

$$\frac{4k^2a_3}{\alpha^2(1-Y)} = \frac{(1-j^2)}{4}$$

where $j$ is an odd integer. Let $a_3 = k_0^2/k^2$, where $k_0$ corresponds to the critical frequency at the maximum electron density. In terms of $k$, this equation becomes

$$\frac{4k_0^2}{\alpha^2[1-(k_1/k)]} = \frac{(1-j^2)}{4}$$

yielding

$$k = k_1 \Big/ \left[1 + \frac{16k_0^2}{\alpha^2(j^2-1)}\right]. \tag{17}$$

It is obvious that $k < k_1$ for all values of $j$, so a discrete set of frequencies, all less than the gyro-frequency, is thereby produced, for which transmission of the circularly-polarized $F$-wave occurs through a perfect ionospheric window. The lowest value of $k$ is given when $j = 3$, namely

$$k = k_1 \Big/ \left(1 + \frac{2k_0^2}{\alpha^2}\right).$$

The reflection coefficient $R_G$, obtained by using the factor $1 + Y$ and not $1 - Y$, can never vanish. This may be seen by noticing that the left-hand side of (11) would be positive, while its right-hand side would be negative.

## 7. Numerical Considerations

If $f_1$ is the gyro-frequency in kilocycles per second, and if $f$ is the general frequency in the same units, then from (17), $R_F$ vanishes when

$$f = \frac{f_1}{1 + [16k_0^2/\alpha^2(j^2-1)]} \tag{18}$$

We shall use the value $f_1 = 1280$ as a realistic figure [this ensures that $Y = 80$ for 16 kc/s, see HEADING and WHIPPLE (1952)].

The distribution of $X$ is given by a two-parameter function, namely

$$X = \left(\frac{k_0^2}{k^2}\right) \operatorname{sech}^2 \frac{\alpha z}{2}.$$

The parameters $k_0$ and $\alpha$ may be chosen so that this distribution fits closely to an actual ionospheric distribution, in spite of irregularities in the latter and possible subsidiary maxima in the lower regions. Fitting an Epstein distribution of electron density to an actual ionosphere (where relevant to do so) is of course an approximate procedure, but the advantage gained is that an analytical solution is available for this model in which the distribution vanishes as $z \to +\infty$. Models



VOL. 20 1961

for which $X \to +\infty$ as $z \to +\infty$, thereby restricting the reflection of long waves to the base of the ionosphere, would not possess the properties here investigated. Here, we shall illustrate the theory using the distribution quoted by ZHEKULIN (1960) (his Fig. 5), based on rocket data.

Now $X = 4\pi Ne^2/mc^2k^2$, where $N$ is the electron density,

so
$$N = \left(\frac{mc^2k_0^2}{4\pi e^2}\right) \operatorname{sech}^2 \frac{\alpha z}{2} . \tag{19}$$

We may calculate $k_0^2$ and $\alpha$ by fitting curve (19) to the physical distribution in various ways.

Firstly, we may identify points on the two curves at values of $N$ given by $N_{\max}$ and $\frac{1}{2}N_{\max}$. These are

$$N_{\max} = 11{\cdot}5 \times 10^5 \qquad \text{when} \qquad z = 0$$

$$\tfrac{1}{2}N_{\max} = 5{\cdot}75 \times 10^5 \qquad \text{when} \qquad z = \pm 65 \text{ km} = \pm 65 \times 10^5 \text{ cm}.$$

Then
$$mc^2k_0^2/4\pi e^2 = 11{\cdot}5 \times 10^5$$

and
$$\tfrac{1}{2} = \operatorname{sech}^2 \frac{\alpha 65}{2} \times 10^5$$

yielding $\alpha = 2{\cdot}7 \times 10^{-7}$ and $k_0^2 = 40{\cdot}6 \times 10^{-7}$ with $k_0^2/\alpha^2 = 5{\cdot}5 \times 10^7$.

Secondly, we may identify the maximum value of $N$ in the two distributions and the total electron content in a vertical cylinder through the ionosphere of unit cross-sectional area. Theoretically, this total is given by

$$T = \int_{-\infty}^{\infty} N\, dz = \frac{mc^2k_0^2}{4\pi e^2} \int_{-\infty}^{\infty} \operatorname{sech}^2 \frac{\alpha z}{2}\, dz$$

$$= \frac{mc^2k_0^2}{4\pi e^2} \left[2\alpha^{-1} \tanh \frac{\alpha z}{2}\right]_{-\infty}^{\infty}$$

$$= \frac{mc^2k_0^2}{\pi e^2 \alpha} .$$

Taken with the equation $N_{\max} = mc^2k_0^2/4\pi e^2$, we see that

$$\alpha = 4N_{\max}/T, \qquad k_0^2/\alpha^2 = \pi e^2 T^2/4mc^2 N_{\max}.$$

A count of the total electron content in the ionosphere quoted by ZHEKULIN gives $T = 14 \times 10^{12}$ and $N_{\max} = 11{\cdot}5 \times 10^5$. Hence

$$\alpha = 3{\cdot}28 \times 10^{-7}, \qquad k_0^2/\alpha^2 = 3{\cdot}75 \times 10^7$$

similar to the first method of calculation.

Using the results of the first calculation, we see from (18) that windows occur for the circularly-polarized wave in which $E_x = iE_y$ at the discrete frequencies given by

$$f = \frac{1280}{1 + [88 \times 10^7/(j^2 - 1)]}$$

where $f$ is in kilocycles per second and $j$ is an odd integer. As long as $88 \times 10^7 \geqslant j^2 - 1$, ($j \leqslant 4000$ say), we may write

$$f = 1280(j^2 - 1)/88 \times 10^7.$$

Small values of $j$ yield frequencies too low to be relevant, so if $j$ lies in the range $1 \leqslant j \leqslant 4000$, we have

$$f = 1{\cdot}45 \times 10^{-6} j^2.$$

This formula embraces all reasonable low frequency waves up to 23 kc/s.

## 8. The Reflection Coefficients

We shall now consider the reflection coefficients for this particular circularly-polarized wave for general frequencies. We have from (10)

$$|R_F|^2 = \frac{1 + \cos 2\pi M}{\cosh(4\pi k/\alpha) + \cos 2\pi M}$$

where

$$M^2 = \tfrac{1}{4} + \frac{4k_0^2}{\alpha^2(Y - 1)} \doteq \frac{4k_0^2}{\alpha^2(Y - 1)}.$$

Now for the above model, $4\pi k/\alpha = 9{\cdot}7f$, where $f$ is in kilocycles per second, so if $9{\cdot}7f$ is large enough for us to approximate to the hyperbolic cosine by an exponential, we have

$$|R_F|^2 = \frac{2\cos^2 \pi M}{(e^{4\pi k/\alpha})/2}$$

and

$$|R_F| = 2|\cos \pi M| e^{-2\pi k/\alpha}.$$

The windows (when $\cos \pi M$ vanishes) therefore occur in a frequency range for which $|R_F|$ is in any case very small. Such a phenomenon contrasts with the curves produced by Hines (1953), in which the reflection coefficient is unity apart from zero values at the windows.

As far as the other circularly-polarized wave is concerned, we must use

$$|R_G|^2 = \frac{1 + \cosh 2\pi P}{\cosh(4\pi k/\alpha) + \cosh 2\pi P}$$

where

$$P^2 = -\tfrac{1}{4} + \frac{4k_0^2}{\alpha^2(1 + Y)} \doteq \frac{4k_0^2}{\alpha^2(1 + Y)}$$

unless the frequency is very low indeed. This modulus $|R_G|$ has the value unity at the frequencies considered, since $2\pi P \gg 4\pi k/\alpha$.

Hence, for a plane-polarized incident wave, we conclude from (13) that $R_{xx} = \frac{1}{2}R_G$, and that $|R_{xx}| = |R_{yy}| = 0{\cdot}5$, the windows not making their presence felt because of the strong competition from the persistent reflection of the $G$-field.

## 9. The Effect of a Collision Frequency

The physical ionosphere, of course, includes a collision frequency $\nu$ that cannot be neglected at low frequencies. As far as the above Epstein model is concerned,



VOL. 20 1961

only a constant collision frequency may be inserted if an analytical solution is still desired. This would be rather a sweeping generalization over the whole range of the ionosphere considered.

The factor $1 \pm Y$ must now be replaced by $1 - iZ \pm Y$, where $Z = \nu/\omega$. This has the effect of making the quantity $M$ complex, and the work of EPSTEIN (1930) shows that the reflection coefficient has the value

$$R = \frac{\Gamma(-2ik/\alpha)\Gamma[(1/2) + M + (2ik/\alpha)]\Gamma[(1/2) - M + (2ik/\alpha)]}{\Gamma(2ik/\alpha)\Gamma[(1/2) + M]\Gamma[(1/2) - M]}$$

where

$$M = \sqrt{\left[\frac{1}{4} - \frac{4k_0^2}{\alpha^2(1 - iZ \pm Y)}\right]} = \beta + i\gamma$$

say. No simple form for the modulus of this expression now exists, unlike (5) and (6) where $M$ was either purely real or purely imaginary. $|R|$ cannot now attain zero values, since no real value of $k$ exists for which the denominator has poles. This may be seen easily, because

$$\Gamma[(1/2) + M]\Gamma[(1/2) - M] = \frac{\pi}{\sin\{\pi[(1/2) + M]\}} = \frac{\pi}{\cos \pi M} = \frac{\pi}{\cos[\pi(\beta + i\gamma)]}.$$

Poles occur only when $\pi(\beta + i\gamma) = j\pi/2$, ($j$ an odd integer), which is clearly impossible if $\gamma \neq 0$, that is, if $Z \neq 0$.

It is to be expected that $|R|$ attains a series of minimum values at a discrete set of frequencies, and that $|T|$ attains a series of maximum values not equal to unity.

Finally, it should be remarked that analytical solutions are out of the question for more realistic variations of $X$ and $Z$ with height, for oblique incidence from below and when the earth's magnetic field is not vertical. The questions of perfect windows or of minimum values of $|R|$ and maximum values of $|T|$ are problems that must then be resolved numerically.

## REFERENCES


| | | |
|---|---|---|
| CLEMMOW P. C. and HEADING J. | 1954 | *Proc. Camb. Phil. Soc.* **50,** 319. |
| EPSTEIN P. S. | 1930 | *Proc. Nat. Acad. Sci., Wash.* **16,** 627. |
| HEADING J. and WHIPPLE R. T. P. | 1952 | *Phil. Trans.* A **244,** 469. |
| HINES C. O. | 1953 | *Quart. Appl. Math.* **11,** 9. |
| NASA SPACE SCIENCE PROGRAM | 1960 | *Proc. I.R.E.* **48,** 438. |
| ZHEKULIN L. | 1960 | *Planet. Space Sci.* **2,** 110. |

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 40 to 48. Pergamon Press Ltd. Printed in Northern Ireland

# Observations on Čerenkov radiation from the night sky

J. White, N. A. Porter and C. D. Long
University College, Dublin

(*Received* 30 *August* 1960)

**Abstract**—Light pulses from extensive air showers have been studied with a system of four wide-angle detectors in coincidence. The zenith angle distribution has been determined by a timing method, and the rate of showers observed under cloud. Most of the light observed appears to originate from altitudes of less than 1 km. The results can be reconciled with previous work if it is assumed that the angular distribution of the light at emission is very wide; the observed characteristics of the shower are then determined principally by the field of view of the receiver.

## 1. Introduction

In principle, the detection of extensive air showers by observing Čerenkov pulses from the night sky (Galbraith and Jelley, 1953) affords a means of studying the early development of showers in the upper atmosphere. It is known from measurements on stars that the absorption of light in the atmosphere is small, and most of the Čerenkov light produced by a shower at all altitudes should reach the ground. Measurements of charged particles, on the other hand, and particularly of electrons, are complicated by the strong absorption in the atmosphere, and the extrapolation from sea-level to the early stages of development is very uncertain.

The relative intensities of the light from different heights observed by a detector on the ground will depend, however, not only on the true absorption in the atmosphere, but also on the angular spread of the light at emission, which is determined by the Coulomb scattering of the particles. If scattering is negligible, all the light from the shower will be concentrated in a circle of about 120 m radius, set by the Čerenkov angle of emission. If Coulomb scattering is large, the light from great heights may be spread over a very large area. The relative contribution from different heights will then be determined largely by the angular field of view of the detector, and by the position and direction of individual showers.

Previous measurements have shown somewhat conflicting results. Čudakov and Nesterova (1958) have found that light can be detected from great heights, and the work of Galbraith and Jelley on the angular decoherence of signals in two detectors, and on the zenith angle dependence (Galbraith and Jelley, 1955 a,b) implies that the high-altitude component is comparable with that from low levels. Both these groups worked with concave mirrors, having a small field of view, and selected shower directions geometrically. The work of Brennan *et al.* (1958) indicates that there is no marked correlation between the direction of the light and the direction of the shower, so that direction selection by geometrical methods may be unreliable.

In the present work, the zenith angle dependence was measured with a timing method, which is not subject to the uncertainty of geometrical selection, and rates of showers were also measured under cloud, in an attempt to determine directly

the fraction of the light received from different altitudes. Measurements on the general characteristics of the background light, and the fraction of time available in which observation was possible, were also carried out. The work was done at sea level, on the Merville site of University College, a suburban area with background conditions intermediate between those in Sydney and at the Pic du Midi.

## 2. Experimental Arrangement

Four detectors were used, of the simple type introduced by Barclay and Jelley (1956). Each consisted of an E.M.I. (Type 6099) 5 in. photomultiplier, mounted in a domestic dustbin (Fig. 1). The field of view with full visibility was 1·7 sterad, very much larger than in previous experiments; light was detected with

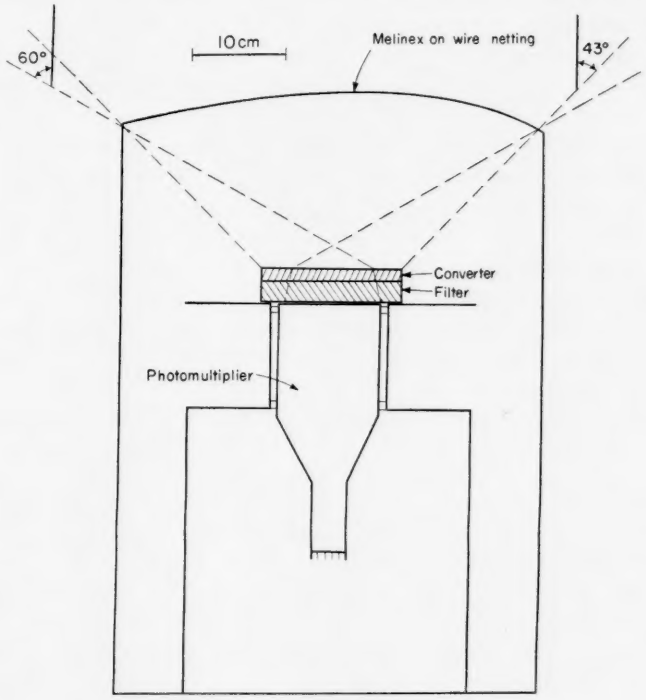


Fig. 1. Detector.

full visibility to an angle of 43° from the zenith, and with reduced visibility to a cut-off at 60°. The sensitive area was 100 cm². Preliminary measurements of photomultiplier currents with different optical filters showed that most of the background light had wavelengths greater than 5000 Å, and blue-violet filters were therefore used throughout the experiment. In addition, ultra-violet converters, either powdered anthracene or dilute amino-G-acid (Saito and Suga, 1959), were placed over the filters in thin polythene containers. The lower wavelength cut-off was set at 3100 Å by Melinex covers on the detectors. The wavelength range accepted was therefore from 3100 to 4800 Å with a mean at about 4000 Å.

The detectors were initially placed close together, and then moved apart to the corners of a square of 80 m side, with vertices roughly N, S, E and W. Balanced signals were fed directly from the collector and last dynode of each multiplier with 300 Ω twin line cable, the only other connexion to the detector being an e.h.t. supply to the photocathode. To reduce misting, the photomultiplier and filters were coated periodically with a commercial preparation, and silica gel was placed in the dustbins. This arrangement operated satisfactorily and heaters were unnecessary.

At the home station, signals from each detector were converted from balance to unbalance with a secondary emission pentode circuit, amplified by a factor 10 at 15 Mc/s bandwidth, and then used to trigger a low-level discriminator. The threshold of the discriminator was set by a ratemeter circuit which operated on the discriminator output, so that the rates of pulses were locked to a preset level. The rate could be varied from 10 to 500 p.p.s., and was normally operated at 10 p.p.s. The discriminator outputs, after an adjustable delay, operated a master circuit, and were also fed to time-amplitude circuits, which converted time delays from adjacent circuits into potentials. These were displayed on the $X$ and $Y$ plates of two cathode ray tubes, which were brightened by a master pulse, and photographed. The positions of the two spots on the tubes correspond to a polar plot of the direction of the shower, assuming that the disc of Čerenkov light is normal to the shower axis; and provided all four detectors were operated by a shower, two semi-independent determinations of shower direction could be made.* Master pulses could be generated by a one-, two-, three- or four-fold coincidence and spots corresponding to each type fell in identifiable groups. The e.h.t. voltages, collector currents and pulse rates were also recorded at half-hour intervals. The system was calibrated with a signal generator at the beginning of each run, and a spark gap at the centre of the square was also used to produce a calibration light pulse, and to measure electronic jitter, which had a maximum value of about 50 nsec. This corresponds to an error of about 10° in measurement near the zenith.

Operation was semi-automatic, the experiment being switched on by hand each night and switched off automatically at a preset time. Sky conditions were monitored by simply photographing the sky with a 35 mm camera, using half-hour exposures, throughout the night. From the number and character of star tracks visible, conditions could be estimated. The method was sensitive even to very slight ground haze. In addition to this, the continuous cloud base measurements from Dublin Airport, about 9 miles away, were available. The agreement between the two records was good.

## Results

### (a) *Sky conditions*

The array was operated for fifty-six nights between 1 January and 1 May 1960, out of a possible sixty nights allowed by moon conditions. Records were obtained for 328 hours of observations. Combining the night sky film with the Collinstown Airport records, clear sky conditions held for 40 per cent of the operating time, clouded for 40 per cent, and hazy for 20 per cent. The overall clear sky time

* The time-amplitude circuit was based on a system designed by T. E. Cranshaw.



VOL. 20 1961

available was therefore 130 hours, or about 5 per cent of the total time (including daylight hours). A reasonable average for the whole year for this climate would probably be about 2 per cent of the total time.

(b) *Background light*

The intensity and spectral distribution of the background were obtained by measuring photomultiplier currents with nine different optical filters, whose characteristics had been determined with a Beckman spectrophotometer. Using the manufacturer's figures for photomultiplier gain, the total intensity between 4000 and 6000 Å was found to be about $2 \times 10^8$ photons/cm$^2$ sterad, about three times the figures found for good viewing conditions (Babcock and Johnson, 1941; Čuveyeev, 1952). The spectral distribution was determined for the four intervals of 1000 Å between 3000 and 7000 Å by solving the nine simultaneous equations in four variables obtained from the filter measurements, and using the S-11 spectral response curve for the photomultiplier. Relative intensities under clear sky conditions were: 3000–4000 Å, 5 per cent; 4000–5000 Å, 26 per cent; 5000–6000 Å, 46 per cent; 6000–7000 Å, 23 per cent. These figures were not accurate to better than 50 per cent in any of the 1000 Å intervals. The distribution differs considerably from that determined spectroscopically for night sky and starlight emission. Under cloud, the total background rose by a factor up to 10, with very strong enhancement of the 5000–6000 Å components, which is presumably caused by sodium and tungsten street lighting. This variation was reduced very considerably when the violet filters were used, however, and on most nights the increase in photomultiplier current under cloud was less than 50 per cent.

During the first 3 months of the experiment the sensitivity controls were set so that the rates of random coincidences were negligible. For the latter 2 months, however, the sensitivity was increased in an attempt to determine the extent to which the fluctuations in background light observed by each detector were correlated in time. The degree of time correlation would be important in any attempt to utilize the methods of Brown and Twiss (1957) to detect Čerenkov radiation. Results from this period were inconclusive, because of increased jitter at high sensitivities, and were not used for the zenith angle and cloud base measurements.

(c) *Variation of rate with cloud base*

For this analysis, nights on which the background was abnormally high were rejected, since the threshold sensitivity of the system was related to the noise level. Only 100 hours of observation were available with background variation of less than 20 per cent, and cloud, when it occurred, was always between 300 and 800 m altitude. The results are shown in Table 1.

Assuming that the distribution of pulse heights from showers has an integral exponent 1·5 (Galbraith and Jelley, 1955 a,b), these figures imply that only about 20 per cent of the light observed at sea level originates above 800 m. The variation in sensitivity with background would tend to increase rather than reduce the apparent contribution from great altitudes, so this figure is probably somewhat over-estimated.

(d) *Zenith angle distribution*

The zenith angle distribution has been determined for a sample of eighty-six three-fold showers. About half of these were also four-fold coincidences. Where a single shower was identifiable on both oscilloscopes the agreement in direction was

Table 1

| Conditions and cloudbase height | Observing time (hr) | No. of showers | Rate/hr |
|---|---|---|---|
| Clear | 30·75 | 57 | 1·85 ± 0·25 |
| 300–800 m | 29·0 | 39 | 1·35 ± 0·21 |
| Ground haze | 23·75 | 17 | 0·72 ± 0·17 |
| Rain | 10·0 | 2 | 0·2 ± 0·14 |
| Snow | 6·5 | 0 | 0 |

good. The distribution is plotted as a histogram in Fig. 2. The smooth curve is the distribution for charged particles in showers found by the M.I.T. Group, (CLARK *et al.*, 1957) and normalized so as to give the same area as the histogram up

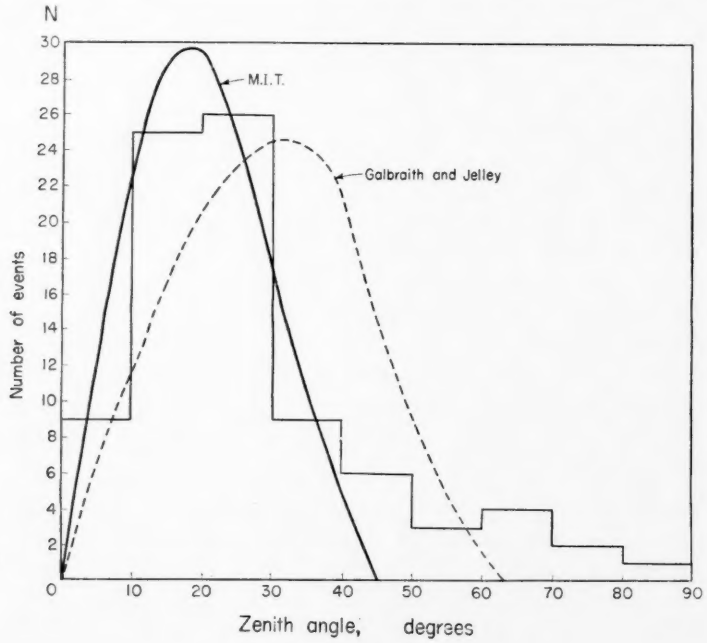


Fig. 2.

to 50°. The dotted curve is the distribution found for light pulses with geometrical selection by GALBRAITH and JELLEY (1955a) (Corrected at angles greater than 43° for the field of view in the present experiment). In the region up to 40° the experimental points are a reasonable fit to the M.I.T. particle distribution, and using a



VOL. 20 1961

similar method of analysis would imply an effective absorption length for the light of about 200 g/cm². This is the shortest absorption length which could be found, since it is roughly equal to the absorption length of the particles which produce the light (Cranshaw *et al.*, 1958a) and implies that the light observed is in equilibrium with the particles at sea level. More showers are observed at zenith angles $>50°$ than would be expected from either of the two curves. These could be due to electronic jitter, to the observation of distant showers, or to the effect of delayed light. This point is considered further in the Discussion. One shower from the sample was recorded at an apparent angle of 105°, caused presumably by electronic jitter. The distribution of showers observed under cloud was similar to that found for clear sky, but the statistics were insufficient to determine an absorption length.

(e) *Lateral distribution of light*

Since the amplitudes of the pulses were not recorded, a direct determination of the lateral distribution was not possible. Some indirect information can be obtained from the decoherence effect. The rate of four-fold coincidences with the detectors close together was 48 $\pm$ 9 per hour. At the 80 m separation it was 0·92 $\pm$ 0·30 per hour. These figures imply a distribution somewhat steeper than that predicted by Gol'danskii and Zhdanov (1954) but as Blatt (1949) has pointed out in connexion with the particle distribution, the decoherence rate is insensitive to the form of the distribution.

The minimum energy of the showers detected may be calculated using an analysis similar to that of Barclay and Jelley (1956) though there are uncertainties arising from the very wide field of view and the spectrum of light accepted. The expected minimum shower size for a single detector would be $10^5$ particles at sea level, (conventionally, about $10^{15}$ eV) corresponding to a flux of twenty-three showers per hour on the 80 m square (Cranshaw *et al.*, 1958b). Since at least three detectors must be struck the minimum shower size must be raised, dependent on the lateral distribution. A value of 3–5 $\times$ $10^5$ particles, was deduced from Monte Carlo calculations, using extreme types of lateral distribution. At this shower size, a rate of one to four showers per hour falls on the area of the array. This corresponds well with the observed rate. The rate of spurious coincidences caused by particles striking the photomultipliers or producing Čerenkov radiation or scintillation light in the converters, is negligible.

## Discussion

The cloud base correlation, zenith angle distribution and to a less definite extent the decoherence rate, are consistent with each other, and imply that the properties of the light at sea level are determined largely by the shower close to the ground. This result is apparently inconsistent with previous work.

There are, however, important differences between this and previous experiments:

(a) The detectors are widely separated from each other.
(b) The light observed is restricted to wavelengths 3100–4800 Å.
(c) The field of view of the detectors is very wide.

The separation of the detectors should enhance, rather than diminish the contribution from great heights. The effect of restricting the light to shorter wavelengths is difficult to estimate. Scattering effects of pure air are negligible over the distances of up to 10 km which could be involved. The effects of water droplets in the atmosphere may be appreciable, but it seems unlikely that the change of wavelength from 5000 Å used by other workers to 4000 Å could produce large effects, since haze scattering is not very strongly dependent on wavelength in the visible (VAN DE HULST, 1952).

The effect of field of view may be very important. The results are, in fact, consistent with an extreme model in which light is emitted isotropically over a hemisphere, when the relative intensities from different heights will fall off almost as an inverse square law. In this model, contributions from different heights depend very largely on the fraction of the shower which is intercepted by the acceptance cone of the detectors. A narrow angle detector receives a relatively large contribution from great heights because more of the shower is visible as the height increases. A wide angle detector will see the whole shower at all altitudes above about 100 m.

On this model the shallow zenith angle distribution found by GALBRAITH and JELLEY (1955a) is consistent with the present observations, since their geometrical selecting system would respond at large zenith angles to vertical showers falling some distance from the receiver. The effect of low energy particles, which are frequently neglected in cascade theory since they do not contribute to the development of the shower, may be dominant in determining the angular distribution of the light produced.

If the emission of light is almost isotropic, however, the validity of the timing method for determining zenith angles may itself be questioned, since it depends on the assumption that a disc of light, normal to the axis, arrives at the ground. BASSI *et al.* (1953) have shown that the particles in showers arrive in a disc, from 1 to 3 m thick, and essentially normal to the shower axis. Light produced by particles which strike the ground close to the detector, and by their ancestors, will arrive simultaneously with the particles and with a time spread of the same order, about 10 nsec. Light from collateral branches of the shower, however, may arrive up to 100 nsec later. Provided the detectors are triggered by the light front, and not by later contributions, the zenith angle determination will be correct, but if the total intensity is close to the detection threshold, and the delayed component greater than the prompt component, an incorrect determination may be made. The overall effect expected would be a broadening of the zenith angle distribution. In the present case, the zenith angle distribution found experimentally is close to the narrowest which could be observed, given the known absorption length of the particles, and this suggests that the effect is small, though the observation of a small number of showers at very large zenith angles may be explained in this way.

These results suggest the possibility of combining narrow and wide angle detectors in a single experiment to study shower development. This would require, however, an accurate determination of the axial position and direction of the shower. It seems possible also that useful information could be obtained by determining the time distribution of light in the pulses, since light from different



VOL.
20
1961

parts of a shower should be spread over a detectable time interval. Some evidence for appreciable time delays has been obtained by one of the authors (N. A. P.) working at A.E.R.E., Harwell, with a 200 Mc/s system, fast photomultiplier, and 5° field of view. About 80 per cent of the pulses observed were found to have durations of 15 nsec or less, but the remaining 20 per cent had durations up to 80 nsec (unpublished work).

## Conclusions

Observations on zenith angle distribution, cloud base correlation and decoherence rate of light pulses, using detectors with a wide field of view, indicate that most of the light observed at sea level originates close to the ground. These results can be reconciled with previous observations if it is assumed that the angular distribution of light from showers is very wide. The field of view of the detector is then dominant in determining the observed characteristics of the shower.

*Acknowledgements*—The authors would like to thank Dr. E. Bretscher, Dr. T. E. Cranshaw and Dr. W. Galbraith for making available electronic equipment constructed and used for cosmic ray work at A.E.R.E. Harwell; Dr. J. V. Jelley for discussion of this paper in manuscript; and Dr. M. Doporto, Director, Irish Meteorological Service, for allowing the use of records from Dublin Airport, as well as Professor T. E. Nevin for help and encouragement in the work. One of us (C. D. L.) would like to thank the Agricultural Research Institute for a research scholarship. Part of the equipment used was constructed by Mr. T. McKenna, and some of the computing work carried out by Mrs. S. B. Porter.

*Note added in proof*

We have recently received a report of Soviet work at mountain altitudes and sea level (Čudakov A. E., Nesterova N. M., Zatsepin V. I. and Tukish E. I. 1960. *Proceedings of the Moscow Cosmic Ray Conference*, Vol. II, p. 50. I.U.P.A.P., Moscow). The Soviet group did not measure the zenith angle distribution at sea level, but their measurements of absolute light intensities seem to imply a greater contribution from high altitudes than is reported here. This discrepancy may arise from geometrical differences between the two experiments.

## References


| | | |
|---|---|---|
| Babcock H. W. and Johnson J. J. | 1941 | *Astrophys. J.* **94,** 271. |
| Barclay F. and Jelley J. V. | 1956 | *Proceedings of the Oxford Conference on Extensive Air Showers* (A.E.R.E. Report). |
| Bassi P., Clark G. and Rossi B. | 1953 | *Phys. Rev.* **92,** 441. |
| Blatt J. M. | 1949 | *Phys. Rev.* **75,** 1584. |
| Brennan M. H., Malos J., Millar D. D. and Wallace C. S. | 1958 | *Nuovo Cim.* (Suppl.) **8,** 662. |
| Brown R. H. and Twiss R. Q. | 1957 | *Proc. Roy. Soc.* A **242,** 300. |
| Clark G., Earl J., Kraushaar W., Linsley J., Rossi B. and Scherb F. | 1957 | *Nature, Lond.* **180,** 353, 406. |
| Cranshaw T. E., De Beer J. F., Galbraith W., Hillas A. M., Norris S. and Porter N. A. | 1958a | *Phil. Mag.* **3,** 811. |
| Cranshaw T. E., De Beer J. F., Galbraith W. and Porter N. A. | 1958b | *Phil. Mag.* **3,** 377. |

| | | |
|---|---|---|
| Čudakov A. E. and Nesterova N. M. | 1958 | *Nuovo Cim.* (Suppl.) **8,** 606. |
| Čuveyeev K. | 1952 | *Dokl. Akad. Nauk SSSR* **87,** 551. |
| Galbraith W. and Jelley J. V. | 1953 | *Nature, Lond.* **171,** 349. |
| Galbraith W. and Jelley J. V. | 1955a | *J. Atmosph. Terr. Phys.* **6,** 250. |
| Galbraith W. and Jelley J. V. | 1955b | *J. Atmosph. Terr. Phys.* **6,** 304 |
| Gol'danskii V. I. and Zhdanov G. B. | 1954 | *Zh. Eksper. Teor. Fiz.* **26,** 405. |
| Saito K. and Suga K. | 1959 | *Nuovo Cim.* **11,** 600. |
| Van De Hulst J. C. | 1952 | *The Atmospheres of the Earth and Planets,* Ed. by G. Kuiper. Chicago University Press. |


VOL.
20
1961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 49 to 58. Pergamon Press Ltd. Printed in Northern Ireland

# The formation of the sporadic-*E* layer in the temperate zones

J. D. WHITEHEAD
Radio Research Board Laboratory, Electrical Engineering Department,
University of Sydney, Australia

(*Received* 2 *September* 1960)

**Abstract**—A horizontal wind shear leads to vertical movement of ions and electrons, which gives rise to the formation of thin layers of ionization. The probability of the layer appearing depends on the horizontal component of the earth's magnetic field. This is the same as the probability observed for sporadic-*E* critical frequency ($f_0E_s$) to exceed 5 Mc/s in different parts of the world. The hypothesis that all temperate zone sporadic-*E* is due to horizontal wind shears is given further support because it shows how sporadic-*E* once formed will tend to persist, gives the correct height at which the layer is most likely to form and is satisfactory so far as orders of magnitude of the various quantities are concerned.

## 1. INTRODUCTION

IN A RECENT communication HEISLER and WHITEHEAD (1960) noted that the fraction of time the critical frequency of the sporadic-*E* layer ($f_0E_s$) exceeded 5 Mc/s depended on the horizontal component of the earth's magnetic field for stations situated in the temperate zone. There was no correlation with total intensity of the magnetic field. The fraction of time increased as the horizontal component increased. The purpose of this paper is to show that this relationship is consistent with sporadic-*E* formation arising from a vertical gradient of horizontal wind. It is shown further that, once formed, such a sporadic-*E* layer tends to persist after the original conditions producing it have changed.

## 2. DERIVATION OF THE VERTICAL VELOCITY OF THE IONS

A horizontal movement of the neutral air in the presence of ions and electrons will tend to move the ions and electrons in varying degrees, polarizing the medium and producing electric fields inside the medium. Further, ions and electrons moving in the earth's magnetic field are acted on by forces at right angles to both the velocity and field. The equations of motion are easily written down and may be solved provided suitable boundary conditions are inserted.

Recently measurements of the velocity of neutral air movements have been made in the lower ionosphere by following meteor trails visually and with radar (MILLMAN, 1959; GREENHOW and NEUFELD, 1959). These measurements have shown that:

(a) the vertical velocity is always small;
(b) the horizontal velocity is the same over horizontal distances exceeding 100 km;
(c) the horizontal velocity often reverses in direction over a height range of about 6 km and has an r.m.s. value of about 25 m/sec.

Although these measurements refer to 80–100 km heights, it seems reasonable to assume that the same general conditions will hold in the 100–120 km height range.

Let us take axes $x$, $y$ and $z$, $z$ being vertically upwards and $x$ being parallel to the horizontal component of the earth's magnetic field.* Let $U_x$, $U_y$ and $U_z$ be the components of the velocity of the air movement. It seems reasonable to assume then that $U_z = 0$ and that $U_x$ and $U_y$ are functions of the height $z$ only. As the horizontal distances involved are about 100 km, we also assume a flat earth and constant direction and magnitude of the earth's magnetic field.

The polarization electric fields will have components $E_x$, $E_y$ and $E_z$. Because of the uniformity over horizontal distance, we will assume that $E_x$ and $E_y$ are constant over the whole region, although $E_z$ is a function of $z$.

Three other conditions are also assumed:

(1) there is no vertical current flow;

(2) the total current flowing in the $x$ direction is zero, although currents in different directions may flow at different heights;

(3) the total current flowing in the $y$ direction is zero.

It is worth noting at this point the order of magnitude of the various quantities involved in the basic equations. At a height of 120 km the electronic collision frequency is about $2 \times 10^4$ per sec. This is between the value given by NICOLET (1953) of $10^4$ per sec and those deduced from the values given by WHITEHEAD (1959a,b) for a height of 110 km ($1{\cdot}3 \times 10^4$ per sec and $1{\cdot}9 \times 10^4$ per sec). The positive ions present are principally those of atomic oxygen having a mass about $3 \times 10^4$ of the electrons. Their angular gyro-frequency is therefore about $3 \times 10^2$ per sec. The electrons have a mean free path about $4\sqrt{2}$ times that of the ions—a little less than this if we take into account the persistence of ionic velocity after each collision. The electron velocity is about $\sqrt{(3 \times 10^4)}$ times the ion velocity. Therefore, the collision frequency of ions is about 1/30 that of electrons or at 120 km about $7 \times 10^2$ per sec—a little greater than the gyro-frequency. As sporadic-$E$ most often occurs below 120 km (for instance, BECKER (1958) gave values of $h'E_s$, 75 per cent of which were equal to or less than 120 km) it appears that the ionic collision frequency is a few times the gyro-frequency in the sporadic-$E$ layer.

Not until we descend below the main part of $E$-region does the electronic collision frequency become equal to the electronic gyro-frequency. We will assume, therefore, that, in the region we are concerned with here, the electrons gyrate many times between collisions.

The space charges existing within the region are much smaller than the total charge of electrons present. We also neglect the presence of negative ions (NICOLET and AIKEN, 1960). Therefore, we can assume (except as regards electrostatic polarization) that the electron density equals the ion density. We further assume that there is only one species of ion.

Starting with the basic equations expressing the fact that the acceleration of the charged particles is zero on the average and inserting the boundary conditions assumed for the problem, it is possible to calculate the (equal) vertical velocity of the electrons and ions. The mathematics is outlined in the Appendix: some care has to be taken in the derivation, because the component of the electric field

* Left handed axes have been used. To convert to right handed axes, change the sign of each $y$ velocity.

parallel to the magnetic field is very small compared to that perpendicular to the magnetic field. It cannot be neglected, however, because it has an appreciable effect on the electron motion.

The expression for the vertical velocity is the following rather complicated expression:

$$
\begin{aligned}
V_z = {} & \frac{\nu_i \omega_{Hi} \cos\theta}{(\nu_i^2 + \omega_{Hi}^2)} \\
& \times \left\{ \left[ \frac{\omega_{Hi}}{\nu_i} (\bar{U}_{yF} - \bar{U}_{yG}) - \sin\theta (\bar{U}_{xF} - \bar{U}_{xG}) \right] \int_{-\infty}^{\infty} G\, dz \int_{-\infty}^{\infty} F\, dz + \right. \\
& + \left[ \frac{\omega_{Hi} \sin\theta}{\nu_i} (U_x - \bar{U}_{xF}) - (U_y - \bar{U}_{yF}) \right] \left[ \int_{-\infty}^{\infty} F\, dz \right]^2 + \\
& \left. + \left[ \frac{\omega_{Hi} \sin\theta}{\nu_i} (U_x - \bar{U}_{xG}) - (U_y - \bar{U}_{yG}) \right] \left[ \int_{-\infty}^{\infty} G\, dz \right]^2 \right\} \div \\
& \div \left\{ \left[ \int_{-\infty}^{\infty} F\, dz \right]^2 + \left[ \int_{-\infty}^{\infty} G\, dz \right]^2 \right\} \qquad (1)
\end{aligned}
$$

where $\nu_i$ = collision frequency of ions
$\omega_{Hi}$ = gyro-frequency of ions
$\theta$ = dip angle
$F = N\nu_i\omega_{Hi}/(\nu_i^2 + \omega_{Hi}^2)$
$G = N\nu_i^2/(\nu_i^2 + \omega_{Hi}^2)$
$N$ = electron density
$\bar{U}_{xF} = \int_{-\infty}^{\infty} F U_x\, dz \Big/ \int_{-\infty}^{\infty} F\, dz$ etc.

### 3. Discussion of the Variation of Vertical Velocity with Height

We first of all see in this expression for $V_z$ that $\bar{U}_{xF}$ etc., are average values of the horizontal velocity taken over a certain height range. For instance, the averages taken with respect to $G$ are roughly mean values of velocity from the base of the ionosphere up to the region where $\nu_i = \omega_{Hi}$, at about 130 km. The averages taken with respect to $F(\bar{U}_{xF}, \bar{U}_{yF})$ are mean values about the region where $\nu_i = \omega_{Hi}$.

In the absence of wind shears (i.e. $U_x$ and $U_y$ independent of height) we see that $V_z = 0$.

This arises because the electrical fields built up by polarization just prevent the vertical movement of ions or electrons.

It is known from the meteor observations that $U_x$ and $U_y$ are subject to quite violent fluctuations over a height range of the order of 10 km. Thus terms such as $U_x - \bar{U}_{xF}$ will generally be larger in magnitude than terms such as $\bar{U}_{xF} - \bar{U}_{xG}$. Furthermore, if we consider the conditions below about 120 km, $\omega_{Hi} < \nu_i$ and to a first approximation we have the following expression for $V_z$

$$V_z \sim \frac{-\omega_{Hi} \cos\theta}{\nu_i} \cdot \Delta U_y \qquad (2)$$

where $\Delta U_y$ is the difference between the actual value of $U_y$ at the height concerned and a mean value between $\bar{U}_{yG}$ and $\bar{U}_{yF}$.

Before this equation can be interpreted in terms of the possibility of sporadic-$E$ formation we must first look at the effects of the vertical velocity of electrons and ions on the electron density distribution.

## 4. The Effect of a Vertical Velocity of Ionization on the Electron Density Distribution in the Ionosphere

We have seen that the vertical velocity depends on $\Delta U_y$, a quantity which may vary rapidly with height. Suppose we take as a model ionosphere one in which the rate of formation of electrons and ions does not vary with height, and each electron-ion pair exists for $T$ sec before recombining. In the absence of vertical movement of ionization the electron density $N_0$ is given by

$$N_0 = qT$$

where $q$ is the rate of formation of electrons.

We now introduce a vertical velocity of ionization given by

$$V_z = -V_0 \sin 2\pi z/\lambda$$

where $\lambda$ is the vertical "wavelength" of the movement. The ions and electrons will move from a height $z = z_0 + Z$ to $z = z_0$ in the time $T$ where

$$TV_0 = \left[\frac{\lambda}{2\pi}\right]\left[\ln \tan \frac{\pi(z_0 + Z)}{\lambda} - \ln \tan \frac{\pi z_0}{\lambda}\right].$$

Then after a time longer than $T$ the electron density at $z = z_0$ is

$$N = qZ/|V_z|.$$

It is now possible to calculate the electron density variation with height. However, our interest centres on the maximum electron density which occurs at $z = 0, \lambda, 2\lambda$, etc. The maximum electron density increase is given by

$$\frac{N_{\max}}{N_0} = \left[\frac{\lambda}{2\pi V_0 T}\right]\left[\exp\left(\frac{2\pi TV_0}{\lambda}\right) - 1\right]. \tag{4}$$

If we take $T = 600$ sec and $\lambda = 10$ km (the random velocity at meteor heights reverses over a vertical distance of about 6 km), the values of maximum electron density increase for values of $V_0$ are given in Table 1.

Table 1. Maximum electron density increases for values of peak vertical velocity of electrons

| $V_0$ (m/sec) | $N_{\max}/N_0$ |
|---|---|
| 3 | 1·7 |
| 6 | 3·2 |
| 9 | 7·4 |
| 12 | 13·4 |
| 15 | 29·5 |



VOL. 20 1961

Thus it is seen that vertical velocities greater than about 6 m/sec lead to very large increases in maximum electron densities.

The meteor observations indicate that $\Delta U_y \sim 25$ m/sec and at a height of 110 km $\omega_{Hi} \cos\theta/\nu_i \sim \frac{1}{2} \cos\theta \sim \frac{1}{4}$. Hence $V_0$ may be expected to have, on occasion, sufficient magnitude to ensure the formation of a dense ionized layer.

Another and more general way of considering the result is to note that $2\pi V_0/\lambda$ is the maximum gradient of vertical velocity. A large maximum of electron density will occur where the vertical velocity is zero and the vertical gradient is large. Taking $T = 600$ sec Table 2 shows the maximum electron density increase for values of vertical velocity gradient with height.

Table 2. Maximum electron density increases for values of vertical gradient of vertical velocity of electrons

| $dV_z/dz$ (per sec) | $N_{\max}/N_0$ |
|---|---|
| $10^{-3}$ | 1·4 |
| $2 \times 10^{-3}$ | 1·9 |
| $4 \times 10^{-3}$ | 4·2 |
| $6 \times 10^{-3}$ | 9·9 |

These electron density increases occur within a time of the order of $T$ sec. If the vertical gradients are suddenly imposed at a time $t = 0$, then after a time $t < T$

$$\frac{N_{\max}}{N_0} = \left[\frac{\lambda}{2\pi V_0 t}\right]\left[\exp\left(\frac{2\pi t V_0}{\lambda}\right) - 1\right]. \tag{5}$$

Thus the electron density will first increase linearly then more rapidly up to the peak value. For wind gradients with a lifetime less than $T$ (small scale turbulence for instance) the net effect is correspondingly reduced. As the time $T$ is really only a mean half-life (5) does not truly represent the growth of a sporadic-$E$ cloud, but we would nevertheless expect a slow growth at the start followed by a rapid increase in critical frequency just before it reaches its maximum. This cannot be observed at one station as we are then unable to distinguish between movement of a cloud of sporadic-$E$ already formed and the growth of the cloud. Experiments are planned at this laboratory to follow the movement and growth of sporadic-$E$ patches using spaced ionosonde recorders.

## 5. The Formation of Sporadic-$E$

We have seen that sporadic-$E$ will form where

(a) the vertical electron (or ion) velocity is zero and

(b) the gradient of vertical velocity is large.

Now the vertical velocity of ions depends on several factors, in particular on

$$\frac{\nu_i \omega_{Hi} \cos\theta}{\nu_i^2 + \omega_{Hi}^2} \tag{6}$$

and on a complicated function involving the horizontal velocity of the neutral air. If the horizontal winds change more rapidly with height than do the functions depending on $\nu_i$, then we would expect sporadic-$E$ formation to occur where the function (6) multiplied by the vertical gradient of horizontal wind was large and electron density was appreciable. Now the vertical gradient of horizontal winds must decrease with height owing to the increase in viscosity, probably as rapidly as $\nu_i$ decreases. Therefore, the gradient of the vertical velocity (of ions) will increase as $\nu_i$ increases, whence sporadic-$E$ will tend to form at the base of the ionosphere at a height where $\nu_i > \omega_{Hi}$. At this height the vertical ion velocity is given by (2).

Let us consider the probability of formation of sporadic-$E$ in different parts of the world, and at different seasons. Summer formation is more likely than winter, because there are more electrons present to be pushed together. Similarly, daytime is more favourable than night-time. (However, it is not suggested that all the seasonal and diurnal variations are accounted for by changes in electron density. Changes in wind gradient must have a large effect.) At different latitudes at midday in summer the base of the ionosphere (taken, say, at the maximum electron density of the $E$-region) varies only a few kilometres in height in the temperate zone. Thus $\nu_i$ at the height at which the layer forms shows little latitude variation. The factor likely to exhibit the most variation over the earth's surface is $\omega_{Hi} \cos \theta$ which is just proportional to the *horizontal* component of the earth's magnetic field. Thus the feature of this type of sporadic-$E$ formation is that it depends on the horizontal component of the earth's field as is found for the fraction of time $f_0E_s > 5$ Mc/s (HEISLER and WHITEHEAD, 1960).

VOL. 20 1961

## 6. THE PERSISTENCE OF SPORADIC-$E$

Sporadic-$E$ observed at one station is found to persist for an average period of 1 hr or so (RAWER, 1958). If it were possible to follow a cloud of sporadic-$E$ easily it might well be found to last much longer than this. A cloud, however, if left to itself would disappear within a few minutes by molecular and eddy diffusion. Continuous formation must occur.

It would appear at first sight that the mechanism proposed in this paper is unsatisfactory because a slight change in horizontal wind at the height of the layer would lead to a layer being formed elsewhere and the original layer quickly disappearing. However (1) suggests that once a thin layer is formed it will tend to continue formation (and hence avoid disappearance) even though the wind configuration leading to the original formation changes.

This comes about in this way: the functions $F$ and $G$ contain $N$, the electron density. If $N$ has a very sharp maximum at one height (the height of the sporadic-$E$ itself) so will $F$ and $G$. Then the values of $\bar{U}_{yG}$, $\bar{U}_{yF}$ etc., will become nearly equal to $U_y$ etc., at this height and $V_z$ will continue to be zero at the height of the sporadic-$E$ layer, although the horizontal winds undergo large changes.

There is no reason why sporadic-$E$ should not, on occasion, form into two or more thin layers as observed by BRIGGS (1958), but it would appear that only one of these layers would persist for more than a few minutes. Indeed, we would expect layer formation to occur at vertical heights separated by several kilometres, this being the vertical semi-wavelength of the horizontal wind system.

It is interesting to speculate here on the possible formation of sporadic-*E* by meteors, by artificial ionization and by turbulence. A weak type of sporadic-*E* is thought to be produced by meteor ionization (NAISMITH, 1954), though one difficulty in explaining its presence has been that such ionization would quickly disappear. However, the mechanism proposed here may explain also how ionization produced by meteors to form a thin layer leads to vertical movements of ionization already present in such a manner as to prolong the life of this layer.

If the air were artificially ionized at the base of *E*-region, a similar mechanism may lead to sporadic-*E* formation. This could be a particularly good test of the theory outlined here if the growth and movement of the cloud of ionization could be followed in some detail.

Another method by which sporadic-*E* could be produced would be by increasing the turbulent velocities. Even if this could only be done artificially in a limited region, there is every possibility that the persistence mechanism could lead to the growth of a large sporadic-*E* cloud.

Recently DUNGEY (1959) has considered the effect of a wind shear in producing irregularities in electron density but neither the dependence on horizontal component of magnetic field nor a means by which a layer once formed might persist is suggested by his results.

## 7. CONCLUSIONS

It appears that horizontal wind shears can satisfactorily explain the formation of sporadic-*E* as regards its height of formation, its global distribution in the temperate zones with respect to the horizontal component of the earth's magnetic field and its long persistence. Measurement of the actual wind shears existing in *E*-region are required before the theory can be checked in detail. Artificial ionization and artificial production of wind gradients may be two means of producing a sporadic-*E* cloud.

*Acknowledgements*—This work has been carried out as part of the investigations of this laboratory sponsored by the Radio Research Board of the Commonwealth Scientific and Industrial Research Organization. The author wishes to thank Dr. G. H. MUNRO and Mr. L. H. HEISLER for helpful discussion.

## REFERENCES


| | | |
|---|---|---|
| BECKER W. | 1958 | AGARDograph 34, *Sporadic-E Ionization* p. 59. N.A.T.O. Advisory Group for Aeronautical Research and Development. |
| BRIGGS B. H. | 1958 | AGARDograph 34, *Sporadic-E Ionization*, p. 151. N.A.T.O. Advisory Group for Aeronautical Research and Development. |
| DUNGEY J. W. | 1959 | *J. Geophys. Res.* **64**, 2188. |
| GREENHOW J. S. and NEUFELD E. L. | 1959 | *J. Geophys. Res.* **64**, 2129. |
| HEISLER L. H. and WHITEHEAD J. D. | 1960 | *Nature, Lond.* In press. |
| MILLMAN P. M. | 1959 | *J. Geophys. Res.* **64**, 2122. |
| NAISMITH R. | 1954 | *J. Atmosph. Terr. Phys.* **5**, 73. |
| NICOLET M. | 1953 | *J. Atmosph. Terr. Phys.* **3**, 20. |
| NICOLET M. and AIKIN A. C. | 1960 | *J. Geophys. Res.* **65**, 1469. |

RAWER K. 1958 AGARDograph 34 *Sporadic-E Ionization*, p. 67. N.A.T.O. Advisory Group for Aeronautical Research and Development.
WHITEHEAD J. D. 1959a *J. Atmosph. Terr. Phys.* **16**, 99.
WHITEHEAD J. D. 1959b *J. Atmosph. Terr. Phys.* **16**, 283.

## APPENDIX

*Derivation of the vertical ion density*

The first six equations listed below arise from the fact that the mean acceleration of the electrons and ions can be taken as zero. $U_x$ and $U_y$ are the components of the velocity of vertical air along the horizontal $x$ and $y$ axes, $V_{xi}$, $V_{yi}$, $V_{zi}$ are the components of ion velocity, $V_{xe}$, $V_{ye}$, $V_{ze}$ the components of electron velocity, $B$ is the magnetic field, $\theta$ the angle of dip, $M$ the ion mass, $m$ the electron mass, $e$ the (positive) ion charge. $E_x$, $E_y$ and $E_z$ are components of the electric field. $B$ is confined to the $x$, $z$ plane. Then we have for ions (using m.k.s. units)

$$0 = (U_x - V_{xi})\,\nu_i + \frac{eE_x}{M} - \frac{BeV_{yi}}{M}\sin\theta \tag{7}$$

$$0 = (U_y - V_{yi})\,\nu_i + \frac{eE_y}{M} + \frac{BeV_{xi}}{M}\sin\theta - \frac{BeV_{zi}}{M}\cos\theta \tag{8}$$

$$0 = -V_{zi}\nu_i + \frac{eE_z}{M} + \frac{BeV_{yi}}{M}\cos\theta \tag{9}$$

and for electrons

$$0 = (U_x - V_{xe})\,\nu_e - \frac{eE_x}{m} + \frac{BeV_{ye}}{m}\sin\theta \tag{10}$$

$$0 = (U_y - V_{ye})\,\nu_e - \frac{eE_y}{m} - \frac{BeV_{xe}}{m}\sin\theta + \frac{BeV_{ze}}{m}\cos\theta \tag{11}$$

$$0 = -V_{ze}\nu_e - \frac{eE_z}{m} - \frac{BeV_{ye}}{m}\cos\theta. \tag{12}$$

The vertical velocities of ions and electrons are equal:

$$V_{zi} = V_{ze} = V_z. \tag{13}$$

We also have that the net horizontal currents are zero:

$$\int_{-\infty}^{\infty} N(V_{xi} - V_{xe})\,dz = 0 \tag{14}$$

and

$$\int_{-\infty}^{\infty} N(V_{yi} - V_{ye})\,dz = 0. \tag{15}$$

We also use the assumption that $E_x$ and $E_y$ do not vary with height.

First of all we derive $V_{xi}$, $V_{yi}$ and $V_z$ from (7), (8) and (9). These give the following results:



VOL. 20 1961

For $V_{xi}$

$$V_{xi}\left(\nu_i^2 + \frac{B^2e^2}{M^2}\right) = U_x\left(\nu_i^2 + \frac{B^2e^2}{M^2}\cos^2\theta\right) - \frac{U_y\nu_i Be}{M}\sin\theta + \left(\frac{eE_x}{M\nu_i}\right)\left(\nu_i^2 + \frac{B^2e^2}{M^2}\cos^2\theta\right) - \frac{eE_yBe}{M^2}\sin\theta - \left(\frac{eE_z}{M\nu_i}\right)\left(\frac{B^2e^2}{M^2}\sin\theta\cos\theta\right) \tag{16}$$

For $V_{yi}$

$$V_{yi}\left(\nu_i^2 + \frac{B^2e^2}{M^2}\right) = \frac{U_x\nu_i Be}{M}\sin\theta + U_y\nu_i^2 + \frac{eE_xBe}{M^2}\sin\theta + \frac{eE_y\nu_i}{M} - \frac{eE_zBe}{M^2}\cos\theta \tag{17}$$

and for $V_{zi}$

$$V_{zi}\left(\nu_i^2 + \frac{B^2e^2}{M^2}\right) = \frac{U_xB^2e^2}{M^2}\sin\theta\cos\theta + \frac{U_y\nu_i Be}{M}\cos\theta + \left(\frac{eE_x}{M\nu_i}\right)\left(\frac{B^2e^2}{M^2}\sin\theta\cos\theta\right) + \frac{eE_yBe}{M}\cos\theta + \left(\frac{eE_z}{M\nu_i}\right)\left(\frac{B^2e^2}{M^2}\sin^2\theta\right). \tag{18}$$

The electric field is composed of a part parallel to and a part perpendicular to $B$. As these have different orders of magnitude it is convenient to write

$$E_x = E_{\parallel}\cos\theta + E_{\perp}\sin\theta$$
$$E_z = E_{\parallel}\cos\theta - E_{\perp}\sin\theta$$

where $E_{\parallel}$ is the component of $E$ parallel to $B$ and $E_{\perp}$ is the component of $E$ perpendicular to $B$ in the $x$, $z$ plane.

Then the quality of $V_{ze}$ and $V_{zi}$ leads to the following relation between $E_{\parallel}$ and $E_{\perp}$ if we remember that $\nu_e \ll \frac{Be}{m}$ and assume that $U_x$ and $U_y$ are of the same order of magnitude.

$$\frac{U_x\nu_i^2B^2e^2}{M^2}\sin\theta\cos\theta - \frac{U_y\nu_iB^3e^3}{M^3}\cos\theta + \frac{eE_y\nu_i^2Be}{M^2}\cos\theta + \frac{eE_{\perp}\nu_iB^2e^2}{M^3}\cos\theta - \left(\frac{eE_{\parallel}B^2e^2\sin\theta}{m\nu_eM^2}\right)\left(\nu_i^2 + \frac{B^2e^2}{M^2}\right) = 0. \tag{19}$$

If the horizontal velocities of ions and electrons are expressed in terms of $E_{\perp}$ and $E_{\parallel}$ instead of $E_x$ over $E_z$, it is found that $E_{\parallel}$ has negligible effect on the ions and

$E_\perp$ has negligible effect on the electrons. Using (19), it is possible to substitute for $E_\parallel$, giving on subtraction

$$V_{xi} - V_{xe} = \left(\frac{\nu_i^2}{\nu_i^2 + \omega_{Hi}^2}\right)\left(U_x + \frac{E_y}{B\sin\theta}\right) + \left(\frac{\nu_i\omega_{Hi}}{\nu_i^2 + \omega_{Hi}^2}\right)\left(\frac{E_\perp}{B\sin\theta} - \frac{U_y}{\sin\theta}\right) \tag{20}$$

and

$$V_{yi} - V_{ye} = \left(\frac{\nu_i^2}{\nu_i^2 + \omega_{Hi}}\right)\left(U_y - \frac{E_\perp}{B}\right) + \left(\frac{\nu_i\omega_{Hi}}{\nu_i^2 + \omega_{Hi}^2}\right)\left(U_x\sin\theta + \frac{E_y}{B}\right) \tag{21}$$

where $\omega_{Hi} = \dfrac{Be}{M}$ = angular gyro-frequency of ions.

Putting $$F = N\nu_i\omega_{Hi}/(\nu_i^2 + \omega_{Hi}^2)$$

and $$G = N\nu_i^2/(\nu_i^2 + \omega_{Hi}^2)$$

and using the integral equations (14) and (15), we have

$$\int_{-\infty}^{\infty} G\left(U_x + \frac{E_y}{B\sin\theta}\right) dz = \int_{-\infty}^{\infty} F\left(\frac{U_y}{\sin\theta} - \frac{E_\perp}{B\sin\theta}\right) dz \tag{22}$$

and

$$\int_{-\infty}^{\infty} G\left(\frac{U_y}{\sin\theta} - \frac{E_\perp}{B\sin\theta}\right) dz = -\int_{-\infty}^{\infty} F\left(U_x + \frac{E_y}{B\sin\theta}\right) dz. \tag{23}$$

Now $E_y$ is constant with $z$, and because from (19)

$$E_\perp/E_\parallel \sim M\nu_i/m\nu_e \gg 1.$$

$E_\perp$ is almost constant with height, therefore $E_y$ and $E_\perp$ can be found from (22) and (23). We now have the information required to calculate $V_z$ and the result is given in (1) of the main part of the paper.



VOL.
20
1961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 59 to 71. Pergamon Press Ltd. Printed in Northern Ireland

# Statistical analysis of radio star scintillation*

S. GRUBER
Department of Electrical Engineering and Research Laboratory of Electronics,
Massachusetts Institute of Technology, Cambridge, Massachusetts

(*Received* 2 *September* 1960)

**Abstract**—Radio star scintillation data are analysed by statistical techniques to find the autocorrelation functions and power density spectra of the fluctuations in amplitude and phase. Interpretation is given in terms of a drifting phase-changing screen as postulated by BOOKER, RATCLIFFE and SHINN. It is shown how the drift velocity of the irregularities in ionospheric electron density that lead to the scintillations, as well as the mean square amplitude and phase of the fluctuations, may be obtained from these records.

## 1. INTRODUCTION

BECAUSE the data analysis of the scintillations of a radio star have not been completely exploited, some of the effects that can be brought out by a statistical approach have perhaps not been realized. In the following discussion it is shown how a statistical analysis of segments of radio star scintillation data has revealed the presence of discrete frequency components in the fluctuations, and has aided in the interpretation of the scintillations. Also, a description of the equipment and the experiment is given. The signals to be processed by the technique of random function analysis are investigated, and the applicability of autocorrelation analysis is considered. The errors introduced by the equipment and the signal analysis are noted. The processed data are discussed, and finally, the analysis is extended to large scintillations.

## 2. EXPERIMENTAL CONDITIONS

In the experiment (see Fig. 1), two receivers, spaced apart on the ground on a north–south baseline, were employed. The radio-frequency sections of these receivers were tuned to either 200 Mc/s or 50 Mc/s and were switched simultaneously at both sites from a central processing point. The 50 Mc/s antennas, which were a section of a parabolic cylinder fed by a half-wave dipole, had a theoretical gain of 8 dB. The 200 Mc/s antennas, which were 28 ft parabolic reflectors fed by a dipole and two reflectors, had a theoretical gain of 25 dB. A single local oscillator was employed to ensure phase coherence. The intermediate frequency amplifiers were tuned to 10·5 Mc/s. The outputs of the two i.f. amplifiers were transmitted through coaxial cables to the central processing station where an analogue multiplication was performed at 10·5 Mc/s; this product was passed through a low-pass filter having a bandwidth of $\frac{1}{2}$ c/s. Two such "detectors" were employed. One had a 90° phase-shifting network in one of the i.f. channels. The output of these detectors was recorded on a two-channel pen recorder to produce a visual record as well as on an f.m. magnetic tape recorder to provide a processible record for later analysis.

* This research was supported in part by the U.S. Army (Signal Corps), the U.S. Air Force (Office of Scientific Research, Air Research and Development Command), and the U.S. Navy (Office of Naval Research).

Basically, except for the magnetic recording and the 90° phase shifting, this system is in effect a Ryle phase-switching interferometer as far as the output is concerned.

The radio star in Cygnus was observed near its upper culmination for 10 min at 50 Mc/s and then for 10 min at 200 Mc/s. This was done during the summer months of 1957 at the Round Hill Field Station (41·54°N latitude, 70·931°W

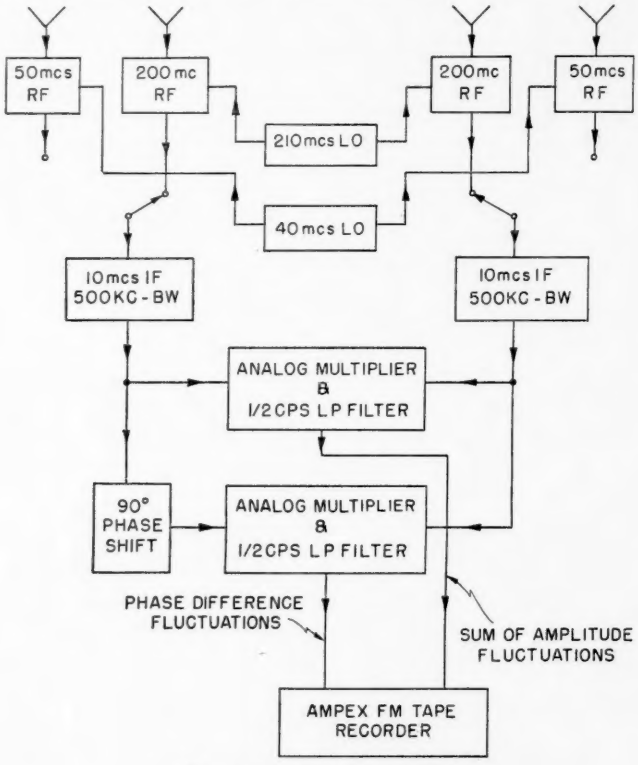


Fig. 1. Scintillation recorder.

longitude) of the Massachusetts Institute of Technology. The receivers were separated by 100, 250 and 400 m in the course of the experiment.

The collected data were sampled three times per second and quantized to sixty-four levels. The digital information was then punched on paper tape in binary form, and autocorrelation functions of the two detector outputs were performed on the Whirlwind I digital computer. Fourier transforms of the auto-correlation functions were derived also by use of the digital computer.

## 3. Signal Analysis

### 3.1. *Nature of the detector output*

As the star signal passes through the ionosphere it becomes modulated in both amplitude and phase. The signal received at a point on earth can be represented as

$$S_A'(t) = [1 + \alpha_A(t)] \exp [-j\phi_A(t)]S_A(t)$$



VOL. 20 1961

where $S_A(t)$ would be the star signal if there were no atmosphere; the fluctuating signals $\alpha_A(t)$ and $\phi_A(t)$ are the slowly varying amplitude and phase modulation and are assumed small compared with unity. The low-frequency term in the product of this signal with the signal received at a point $B$ is approximately

$$\overleftrightarrow{S_A'(t)S_B'^*(t)} = [1 + \alpha_A(t)][1 + \alpha_B(t)] \exp\{-j[\phi_A(t) - \phi_B(t)]\}\overline{S_A(t)S_B^*(t)} \quad (1)$$

where the double arrow denotes that the signal has been filtered sufficiently to leave only the d.c. component of $\overline{S_A(t)S_B^*(t)}$ [the asterisk meaning the complex conjugate of $S_B(t)$] and the low-frequency modulation. It is now assumed that the spacing between the receivers is small enough to cause the r.f. phase difference between the two receivers to be zero (the star is now assumed to be very nearly overhead during the time of observation). Then $\overline{S_A(t)S_B^*(t)}$ is just the star power, $P_s$. If the phase modulation is small, $\phi \ll 1$, the output of the detector with no phase shift is half the real part of (1), that is,

$$\tfrac{1}{2}\mathrm{Re}\left\{\overleftrightarrow{S_A'(t)S_B'^*(t)}\right\} = P_s[1 + \alpha_A(t) + \alpha_B(t)] \quad (2)$$

and the output of the detector, which has a channel with a 90° phase shift, is half the imaginary part of the same expression:

$$\tfrac{1}{2}\mathrm{Im}\left\{\overleftrightarrow{S_A'(t)S_B'(t)}\right\} = P_s[\phi_A(t) - \phi_B(t)]. \quad (3)$$

The foregoing analysis pertains to small scattering wherein multiple scattering is neglected. Hewish (1952) and others have found intense fluctuations in amplitude, and so the data must be examined with this in mind.

Booker *et al.* (1950) show that there is a one-to-one relationship between the spatial autocorrelation function of the index of refraction of a diffracting screen and the pattern of the fluctuations on the ground. Whatever the model is for the medium causing the scintillations, the fluctuations in amplitude, $\alpha(t)$, and in phase, $\phi(t)$, are linearly related to the medium that causes them in the small-scintillation approximation. As shown by Hewish (1952) and others, the pattern of the fluctuations moves past the receiver practically intact; hence it would appear to be of interest to determine the autocorrelation function of the fluctuations to see what information it contains about the medium.

In all of the following, for simplicity of discussion, the Booker *et al.* model will be used; that is, a stable pattern of random intensity drifts with constant velocity, and the autocorrelation function of the ground pattern is identical to that of the diffracting screen. However, most remarks will apply to other models also.

### 3.2. *Autocorrelation analysis of the data*

The autocorrelation function is approximated over the 10 min data run as

$$R_\alpha(\tau) \approx \frac{1}{T}\int_0^T \alpha(t)\alpha(t + \tau)\, dt$$

where $T$ is the run time. Since $R(\tau)$ is negligible for $\tau$ on the order of a fraction of

$T$, this yields a good approximation. The cross-correlation function between receivers $A$ and $B$ is

$$R_{\alpha_{AB}}(\tau) \approx \frac{1}{T}\int_0^T \alpha_A(t)\alpha_B(t+\tau)\,dt.$$

Similar statements hold for the phase fluctuations.

After the d.c. component is removed, the autocorrelation function of the signal at the output of the detector without the phase-shifted channel is given by

$$R_{\text{amp}}(\tau) = R_{\alpha_A}(\tau) + R_{\alpha_B}(\tau) + R_{\alpha_{AB}}(\tau) + R_{\alpha_{BA}}(\tau).$$

If the region is homogeneous, the resulting autocorrelation function is

$$R_{\text{amp}}(\tau) = 2R_{\alpha}(\tau) + R_{\alpha_{AB}}(\tau) + R_{\alpha_{AB}}(-\tau). \tag{4a}$$

Similarly, the autocorrelation function of the channel with the 90° phase shift is

$$R_{\text{phase}}(\tau) = 2R_{\phi}(\tau) - R_{\phi_{AB}}(\tau) - R_{\phi_{AB}}(-\tau). \tag{4b}$$

If the assumption of a drifting diffracting screen is used,

$$R_{\text{amp}}(\tau) = 2\rho_N(v_N\tau)\rho_E(v_E\tau) + \rho_E(v_E\tau)[\rho_N(D - v_N\tau) + \rho_N(D + v_N\tau)] \tag{5a}$$

and

$$R_{\text{phase}}(\tau) = 2\rho_N(v_N\tau)\rho_E(v_E\tau) - \rho_E(v_E\tau)[\rho_N(D - v_N\tau) + \rho_N(D + v_N\tau)] \tag{5b}$$

where $v_N$ is the pattern velocity towards the north, and $v_E$ is the pattern velocity towards the east, and the product $\rho_N(x)\rho_E(y)$ is the spatial autocorrelation function of the diffracting screen. In the foregoing, the fluctuations are assumed to be small. Also it is assumed that the screen may be anisotropic with principal axis in the $x$ (north–south) and $y$ (east–west) directions and that the autocorrelation function can be factored into the product of two functions, one for the spatial autocorrelation function in a north–south direction and the other for the spatial autocorrelation function in the east–west direction. The receivers are spaced apart a distance $D$ in the north–south direction.

When the receiver spacing is small compared with the correlation distance of the diffracting screen in the north–south direction, one can use the first few terms of a Taylor expansion of the spatial autocorrelation function $\rho_N(x)$ about $x = v_N\tau$ to obtain

$$R_{\text{amp}}(\tau) \approx 4\rho_N(v_N\tau)\rho_E(v_E\tau) \tag{6a}$$

and

$$R_{\text{phase}}(\tau) \approx -\rho_E(v_E\tau)D^2 \left.\frac{\partial^2\rho_N(x)}{\partial x^2}\right|_{x=v_N\tau} \tag{6b}$$

3.3. *Power density spectra*

The power density spectrum of a random function is the Fourier transform of its autocorrelation function. Hence, the power per unit bandwidth in the scintillations is

$$G_{\text{amp}}(\omega) = 2\int_{-\infty}^{\infty} \rho_N(v_N\tau)\rho_E(v_E\tau)\exp(-j\omega\tau)\,d\tau + \int_{-\infty}^{\infty} \rho_E(v_E\tau)[\rho_N(D - v_N\tau) + \rho_N(D + v_N\tau)]\exp(-j\omega\tau)\,d\tau.$$



VOL. 20 1961

If we define

$$G_E(\omega) = \int_{-\infty}^{\infty} \rho_E(v_E\tau) \exp(-j\omega\tau)\, d\tau$$

and

$$G_N(\omega) = \int_{-\infty}^{\infty} \rho_N(v_N\tau) \exp(-j\omega\tau)\, d\tau$$

we can write

$$\int_{-\infty}^{\infty} \rho_N(D + v_N\tau) \exp(-j\omega\tau)\, d\tau = \exp\left(+j\frac{\omega D}{v_N}\right) G_N(\omega)$$

and

$$\int_{-\infty}^{\infty} \rho_N(D - v_N\tau) \exp(-j\omega\tau)\, d\tau = \exp\left(-j\frac{\omega D}{v_N}\right) G_N(\omega).$$

The equation for the power density spectrum of the amplitude now can be written:

$$G_{\text{amp}}(\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} G_E(\omega - \xi) G_N(\xi)\left[1 + \cos\xi\frac{D}{v_N}\right] d\xi. \tag{7a}$$

Similarly,

$$G_{\text{phase}}(\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} G_E(\omega - \xi) G_N(\xi)\left[1 - \cos\xi\frac{D}{v_N}\right] d\xi. \tag{7b}$$

In the situation where the receiver spacing is small compared with the north–south correlation distance, we find

$$G_{\text{amp}}(\omega) \approx \frac{2}{\pi}\int_{-\infty}^{\infty} G_E(\omega - \xi) G_N(\xi)\, d\xi \tag{8a}$$

and

$$G_{\text{phase}}(\omega) \approx \frac{D^2}{2\pi v_N{}^2}\int_{-v_{N/D}}^{v_{N/D}} G_E(\omega - \xi)\xi^2 G_N(\xi)\, d\xi. \tag{8b}$$

Since observations indicate that the drift is predominantly in an east–west direction, it is useful to note that for the case of $v_N = 0$, the equations for the spectra become

$$G_{\text{amp}} = \frac{1}{\pi} G_E(\omega)[\rho_N(\text{o}) + \rho_N(D)] \tag{9a}$$

and

$$G_{\text{phase}}(\omega) = \frac{1}{\pi} G_E(\omega)[\rho_N(\text{o}) - \rho_N(D)]. \tag{9b}$$

Equations (9a) and (9b) show that if the phase fluctuations are a result of the same ionospheric inhomogeneities that cause the amplitude fluctuations, and if the drift is mainly in the east–west direction, the phase spectrum should be identical to the amplitude spectrum.

3.4. *Validity of autocorrelation analysis of scintillation data*

Because the intensity of the scintillations varies from day to day as well as nocturnally, it is apparent that the statistics of radio star scintillations are not

stationary; and in order for an autocorrelation analysis, as discussed above, to be valid, the statistics should not be time-dependent. One would expect that within a short period, say during the 20 min of observation, the fluctuations could be considered statistically stationary. Therefore, the conclusions drawn from the statistical analysis of the short data runs should be reasonable. But we cannot draw general conclusions about the diurnal variations in the scintillating region on the basis of our 10 min analysis. If, for example, the phase and amplitude spectra are alike, we should not conclude that the same region is always the cause of both types of fluctuation.

3.5. *Limitations and errors in the measurements*

The sensitivity of the receiving equipment is inversely proportional to the sum of the receiver noise temperature and sky temperature and to the bandwidth of the detector's low pass filter. This means that there is an upper limit on the frequency components of the fluctuations that can be measured. In our case, $\frac{1}{2}$ c/s was the half-power bandwidth of the two-section R-C filter employed. The low-frequency end of the power density spectrum of the scintillations is limited by the condition that it would be preferable to have the lowest frequency component oscillate a number of times during the interval of observation so that the averaging inherent in the autocorrelation function would be meaningful. In our case the low-frequency end of the spectrum was attenuated below approximately 1/60 c/s. This causes the autocorrelation functions of band limited noise to cross the axis at a value of $\tau$ somewhat different from the correct one, but this error is quite small for the records having faster scintillation. This error is typically about 15 per cent for slow fluctuations (see Fig. 2). The effect on the power density spectra is to remove the power below 1/60 c/s. Actually the processed data will show components of the fluctuations down to zero frequency (see Fig. 7), but these components result from the Fourier transform technique.

The Fourier transforms of the autocorrelation functions were performed on a digital computer by means of a program that applied a weighting function technique of Ross (1954) to minimize the effects of finite function length. This causes a slight smearing of the spectrum with the result that a discrete component has a finite line width. Even without this weighting, a specular component would have a finite width because the autocorrelation functions are only approximately 100 sec in length and, therefore, the line width would be on the order of 1/100 c/s. The use of the weighting function reduces to a tolerable level the secondary peaks which occur when a function of finite time duration is transformed, but the smearing of the curve is somewhat greater than the 1/100 c/s width mentioned before.

## 4. Discussion of Data

The processed data are displayed in the form of a graph on an oscilloscope screen, and photographs of the presentation are taken. The plots in the left-hand columns of Figs. 2–7 are the autocorrelation functions (normalized to unity at $\tau = 0$), each mark on the horizontal axis indicating intervals of 8/3 sec. The power density spectra are in the right column opposite the corresponding autocorrelation functions. Each mark on the horizontal axis represents 1/50 c/s.



VOL. 20 1961

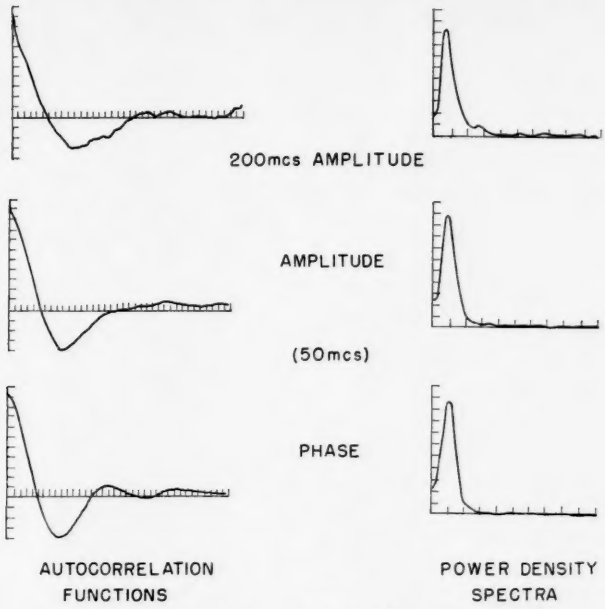


Fig. 2. Typical record of slow fluctuations.

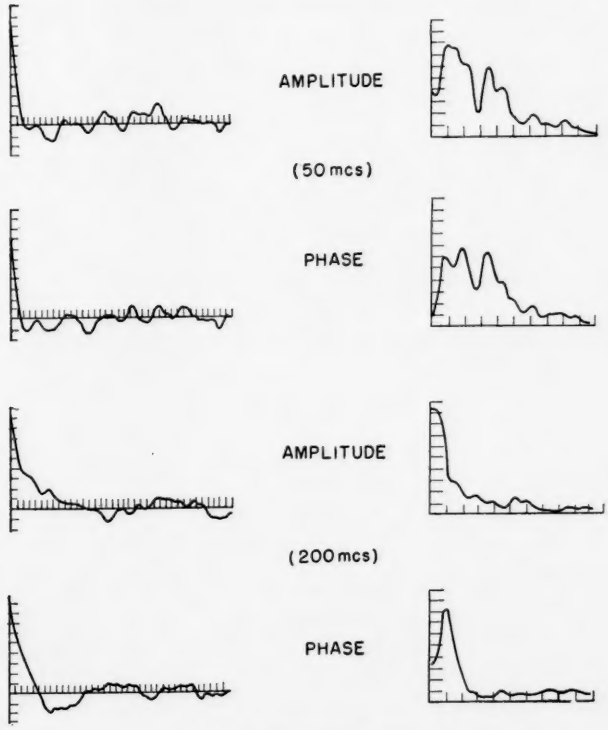


Fig. 3. 20 August 1957.



5—(20 pp.)

We find that the data, in general, indicate that the phase and amplitude have similar spectra. This leads to several conclusions:

1. The assumption of an east–west drift velocity and a north–south elongatio.

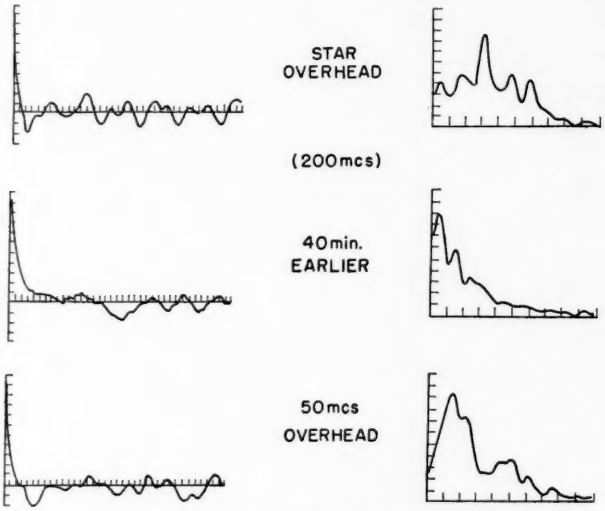


Fig. 4. Amplitude records, 3 September 1957.

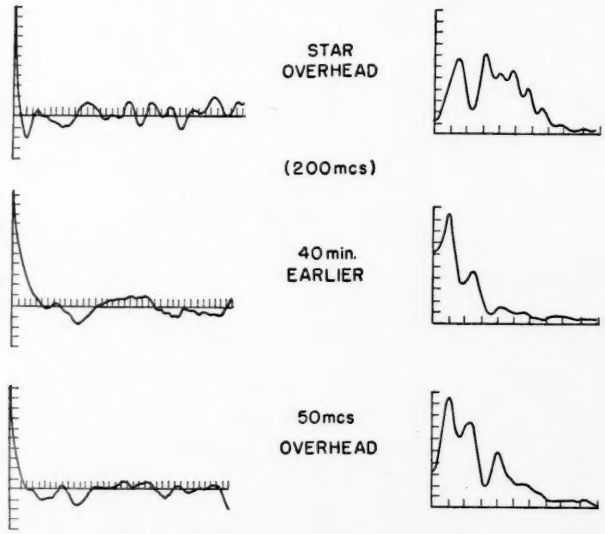


Fig. 5. Phase records, 3 September 1957.

of the irregularities is not contradicted by the data. (Since this experiment did not measure the drift velocity, the possibility of the north–south correlation length being smaller than the receiver spacing cannot be ruled out by these results.)

2. The region causing the amplitude fluctuations is probably the same one



VOL.
20
1961

that introduces the phase fluctuations since it is difficult to conceive of a way in which the phase and amplitude scintillations could be introduced at different heights of the ionosphere and yet appear so similar in their power density spectra. This appears to refute an argument by BOOKER (1956) that uses a relationship derived by HEWISH (1952) and assumes the same size of fluctuations throughout the ionosphere; the phase fluctuations reach a maximum at approximately 100 km, whereas the amplitude fluctuations occur mainly at a level of 150 km.

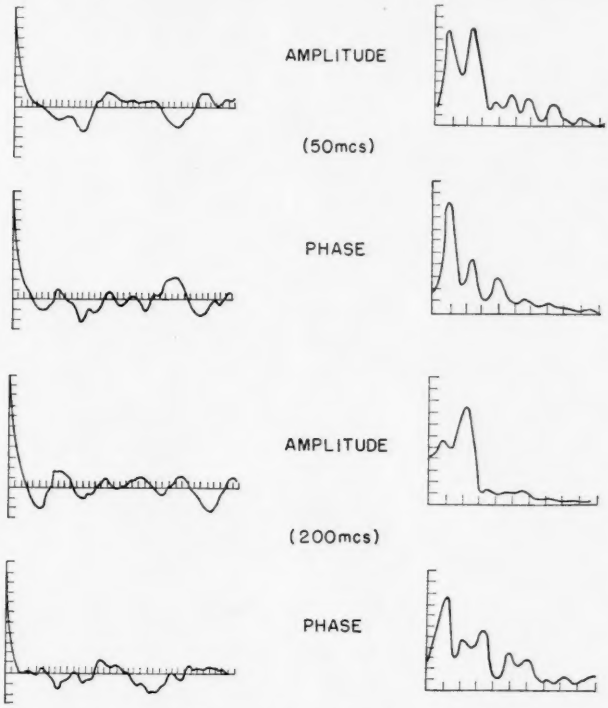


Fig. 6. 4 September 1957.

3. The scintillations at 50 Mc/s have at least as much relative power in the high frequency components as those at 200 Mc/s. Generally, the spectrum of both the 50 Mc/s and 200 Mc/s fluctuations have about the same relative power distribution. This is in agreement with the observations of HEWISH (1952). On occasion, the scintillation at 50 Mc/s extends to noticeably higher frequencies than the scintillations at 200 Mc/s. (As an example, see Fig. 3.) We expect this to be the case when the fluctuations are very intense, as pointed out by HEWISH (1951). In Section 5, there is a simple extension of HEWISH's analysis; an application of the results of the analysis to the record of Fig. 3 is given in the Appendix. The autocorrelation functions for the amplitude fluctuations at 200 Mc/s and 50 Mc/s lead to an estimate of the mean-square phase change through the region causing the scintillations.

4. One of the very striking features of the scintillations is the presence of

discrete frequency components. This is evidenced by the peaks in the spectra and the oscillatory nature of the autocorrelation functions and is most vividly indicated by the 200 Mc/s records shown in Figs. 4 and 5. It can be seen on other records as well (see Fig. 6, for example). It is quite possible that this is the manifestation of the so-called ridge or the ribbed irregularities of WILD and ROBERTS (1956), and that the high-frequency components in the spectra peaks are very fine structures not revealed by their type of presentation. A more specific discussion of the data follows. It is included to illustrate some of the conclusions that result from the autocorrelation function analysis technique.

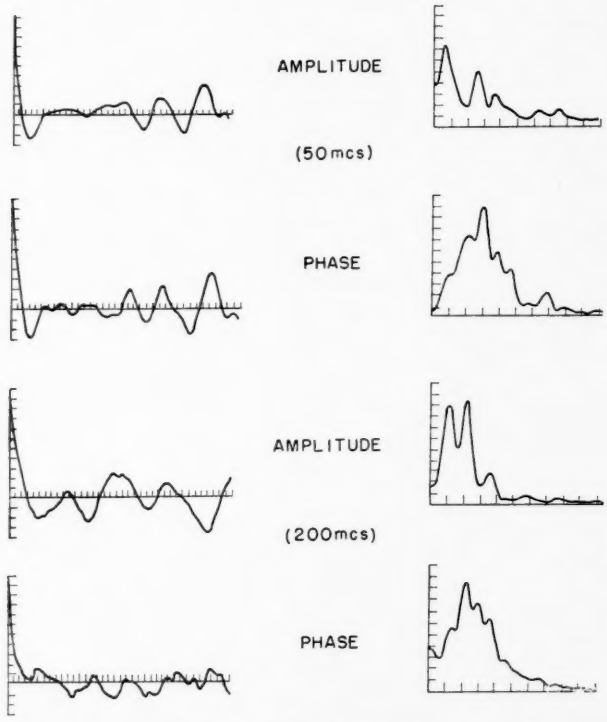


Fig. 7. 14 August 1957.

Fig. 7 shows a marked difference between the amplitude and phase scintillations, but the 50 and 200 Mc/s records are similar in structure. Further, we notice that the phase power spectra have about the same high-frequency drop-off. The same may be said for the amplitude records. These facts lead us to conclude that the scattering is weak and that since the higher frequencies are more emphasized on the phase-difference records, there is a drift velocity component along the receiver base line (north–south). A rough estimate can be made of this velocity, since the antenna separation is known. Equation (8b) essentially states that the phase spectrum, under the approximations stated, is the convolution of the spectrum resulting from the east–west drift with $\omega^2 G_N(\omega)$. Since $G_E(\omega)$ has its maximum at



VOL.
20
1961

$\omega = 0$ and $\xi^2 G_N(\xi)$ can be expected to have its maximum at the upper limit of the integral $v_N/D$, then the maximum of the integral should occur near $\omega = v_N/D$, provided that the amplitude spectrum reaches a peak at a lower frequency. The phase spectrum at 50 Mc/s reaches a peak at approximately 3/50 c/s, whereas the peak at 200 Mc/s occurs at approximately 1/20 c/s. When the smaller figure is used, the drift is found to be $|v_N| \approx 100$ m/sec for the 400 m antenna separation.

Figs. 4 and 5 show the long-time non-stationarity in the fluctuations. A recording was taken of the scintillations with the source 10° before reaching its zenith. The change in conditions is apparent on the 200 Mc/s records. The fluctuations in the 50 Mc/s records were very strong, as were the fluctuations at 200 Mc/s. However, the character of the 50 Mc/s record changed little, whereas the 200 Mc/s fluctuations had relatively greater emphasis on the higher frequency components during the period near the star's zenith. This fact does not seem to agree with the Booker *et al.* model. A similar observation has been made by Chivers (1960).

## 5. Strong Fluctuations

The Booker *et al.* paper proves a theorem that states that the spatial autocorrelation function just below the phase-changing screen in the ionosphere is identical to the spatial autocorrelation function of the electric field on any plane parallel to the screen if the smallest irregularities in the screen are greater than the wavelength. It is a simple matter to determine the spatial autocorrelation function just below the screen since only phase fluctuations appear. The electric field immediately below the screen at a position in the plane $x$, $y$ is given by

$$E(0, y) \sim A \exp [j\phi(y)]$$

and the field at the position displaced from $y$ by $\xi$ in the $y$ direction is given by

$$E(0, y + \xi) \sim A \exp [j\phi(y + \xi)].$$

The autocorrelation function is the average of the product of all fields measured a distance $\xi$ apart.

$$R(\xi) = \overline{E(0, y)E^*(0, y + \xi)}^{y}$$

where the asterisk denotes that the complex conjugate is taken.

Assuming, as Hewish (1951) does, that the phase fluctuations introduced by the screen have a gaussian distribution, we find that

$$R(\xi) \sim \int_{-\infty}^{\infty} \frac{\exp \{-j[\phi(y) - \phi(y - \xi)]\}}{2\pi\phi^2[1 - \rho(\xi)]} \exp \left\{ - \frac{[\phi(y) - \phi(y - \xi)]^2}{2\phi^2[1 - \rho(\xi)]} \right\} dy$$

$$R(\xi) \sim \exp \{-\phi^2[1 - \rho(\xi)]\}$$

Here $\rho(\xi)$ is the normalized spatial autocorrelation of the phase changes in the screen in the $y$-direction. This function, $R(\xi)$, is the same as the spatial autocorrelation function of the fields on the earth. The remarks of Hewish (1951) on $R(\xi)$ hold, only now we have the explicit form of this function.

The autocorrelation functions of the amplitude fluctuations determined experimentally are $R(\xi = v\tau) - R(\infty)$, normalized at $\xi = 0$. If the fluctuations are assumed to exist only in the east–west direction, then

$$R_{\text{amp}}(\tau, \lambda) = \frac{\exp\{-\phi^2(\lambda)[1 - \rho_E(v_E\tau)]\} - \exp\{-\phi^2(\lambda)\}}{1 - \exp\{-\phi^2(\lambda)\}}$$

where $\lambda$ is the wavelength of the radiation that is under observation. Since the mean-square fluctuations in phase, $\overline{\phi^2}$, vary directly as the square of the wavelength and since the records were taken at two frequencies almost simultaneously, we can determine both $\overline{\phi^2}(\lambda)$ and $\rho_E(v_E\tau)$.

## 6. Conclusions

The simple model of a drifting diffraction grating seems to explain some of the general characteristics of radio star scintillations; however, it appears that the grating is not necessarily random, as usually supposed. This was indicated by Wild and Roberts and is emphasized here by the discrete frequency components in the spectra.

The technique of generalized harmonic analysis of random functions, when applied to radio star scintillations, can make evident the nature of these scintillations. The extent of the phase modulation introduced by a diffracting screen can be seen from the spectra of the fluctuations of two observation frequencies. The validity of the model can also be ascertained by the character of the scintillation spectra. Further, interferometer spacing and spectra can yield drift velocity information.

*Acknowledgements*—The author wishes to acknowledge the constructive criticism of Prof. J. B. Wiesner and also the unselfish efforts of Mr. Arthur E. Berg during the experimental phase.

## References


| | | |
|---|---|---|
| Booker H. G. | 1956 | *J. Geophys. Res.* **61**, 673. |
| Booker H. G., Ratcliffe J. A. and Shinn D. H. | 1950 | *Phil. Trans.* **242**, 579. |
| Chivers H. J. A. | 1960 | *J. Atmosph. Terr. Phys.* **17**, 181. |
| Hewish A. | 1951 | *Proc. Roy. Soc.* A **209**, 81. |
| Hewish A. | 1952 | *Proc. Roy. Soc.* A **214**, 494. |
| Ross D. T. | 1954 | Servomechanisms Laboratory, Report No. 7138-R-5, M.I.T. |
| Wild J. P. and Roberts J. A. | 1956 | *J. Atmosph. Terr. Phys.* **8**, 55. |


## Appendix

From the analysis of Section 5 we can determine $\overline{\phi^2}(\lambda)$ and $\rho_E(v_E\tau)$. Let us consider the data in Fig. 3. For the time $\tau_0 \approx 4$ sec, $R_{\text{amp}}(\tau_0, 50) \approx 0{\cdot}25$ and $R_{\text{amp}}(\tau_0, 200) \approx 0{\cdot}40$.

Let us assume that $\phi^2(200) \ll 1$, with the result that

$$R_{\text{amp}}(\tau_0, 200) \approx \rho(\tau_0) \approx 0{\cdot}4$$



VOL. 20 1961

then by a trial-and-error process we find that

$$\overline{\phi^2}(200) \approx 0{\cdot}068$$

and

$$\overline{\phi^2}(50) \approx 0{\cdot}109$$

Of course, we have assumed here that the model of BOOKER *et al.* applies to the data, at least for small-time shifts.

OL.
20
961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 72 to 84. Pergamon Press Ltd. Printed in Northern Ireland

# RESEARCH NOTES

## International Geophysical Calendar for 1961

(*Received* 26 *November* 1960)

1. *Purpose*

The accompanying International Geophysical Calendar 1961 designates some special days and intervals for special attention for geophysical experiments and analysis. The Calendar serves to encourage world-wide coordination of observation or analysis of those geophysical phenomena which vary significantly during the course of a year. These phenomena are mainly in the scientific disciplines dealing with the earth's atmosphere. In some experiments, such as the routine measuring of variations of the earth's magnetic field, the observing and analysis programs at observatories are carried out at a uniform level throughout the year; in these cases the Calendar is not needed. However, in many other experiments (e.g. rocket experiments) it is not practical or meaningful to carry out the same program on each and every day. Here the Calendar can provide a useful mechanism for coordination. Experimenters will know that their colleagues in other laboratories and in other disciplines will tend to also carry out experiments on the days or intervals marked on the Calendar. In this way, results of experiments may be more easily and usefully compared.

In some scientific fields, international scientific organizations have made specific recommendations for programs to be done on days or intervals marked on the calendar. In others, the arrangements are informal or self-evident. Some examples are given below, along with the criteria for selection of dates.

2. *Regular World Days* (*RWD*)

These are 3 consecutive days each month. They always come in the middle or just after the middle of the month and include where possible the times of equinox and solstice. They come in the middle of the week—Tuesday, Wednesday and Thursday. The groups of RWD are evenly spaced through the year so far as it is practical. They include wherever possible days of solar eclipse and meteor showers. One RWD each month (always Wednesday) is designated the RWD of highest priority.

The RWD are intended for experiments or observational programs or analysis programs which as a practical matter can be done only about 10 per cent of the time and should be spaced through the year. Further it is suggested that whenever there are no special reasons for selecting some other days, the RWD be used for unusual or special experiments. This applies perhaps particularly to scientific fields in atmospheric geophysics, including aspects of cosmic rays, meteorology, airglow, ionosphere, geomagnetism, aeronomy. Examples in Ionospheric Physics are: oblique incidence pulse transmission and reception; absorption measurements by pulse reflection technique; extended observing schedule for whistlers and v.l.f. emissions; vertical sounding observing schedule faster than normal; more detailed reduction of vertical sounding ionograms by $f$-plot, $h'$-plot, etc.; hourly reduction from ionograms of $F$-region true height parameters "$hc$" and "$qc$".



VOL.
20
196

# International Geophysical Calendar 1961

Issued November 1960 by the International World Day Service under the auspices of U. R. S. I.

**1961 JANUARY 1961**

| S | M | T | W | T | F | S |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| 15 | 16 | (17) | (18) | (19) | 20 | 21 |
| * 22 | 23 | 24 | 25 | 26 | 27 | 28 |
| 29 | 30 | 31 | | | | |

**1961 FEBRUARY 1961**

| S | M | T | W | T | F | S |
|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 |
| 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| (12) | 13 | (14) | [15] | (16) | 17 | (18) * * |
| 19 | 20 | 21 | 22 | 23 | 24 | 25 |
| 26 | 27 | 28 | | | | |

**1961 MARCH 1961**

| S | M | T | W | T | F | S |
|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 |
| 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| 19 | 20 | (21) | (22) | (23) | 24 | 25 |
| 26 | 27 | 28 | 29 | 30 | 31 | |

**1961 APRIL 1961**

| S | M | T | W | T | F | S |
|---|---|---|---|---|---|---|
| | | | | | | 1 |
| 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 16 | 17 | (18) | (19) | (20) | 21 | 22 |
| 23 | 24 | 25 | 26 | 27 | 28 | 29 |
| 30 | | | | | | |

**1961 MAY 1961**

| S | M | T | W | T | F | S |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| 14 | 15 | (16) | (17) | (18) | 19 | 20 |
| 21 | 22 | 23 | 24 | 25 | 26 | 27 |
| 28 | 29 | 30 | 31 | | | |

**1961 JUNE 1961**

| S | M | T | W | T | F | S |
|---|---|---|---|---|---|---|
| | | | | 1 | 2 | 3 |
| 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| 18 | 19 | (20) | (21) | (22) | 23 | 24 |
| 25 | 26 | 27 | 28 | 29 | 30 | |

**1961 JULY 1961**

| S | M | T | W | T | F | S |
|---|---|---|---|---|---|---|
| | | | | | | 1 |
| 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| (16) | 17 | (18) | (19) | (20) | 21 | (22) |
| 23 | 24 | 25 | 26 | 27 | 28 | 29 |
| 30 | 31 | | | | | |

**1961 AUGUST 1961**

| S | M | T | W | T | F | S |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 |
| 6 | 7 | 8 | 9 | 10 | [11] | 12 |
| 13 | 14 | (15) | (16) | (17) | 18 | 19 |
| 20 | 21 | 22 | 23 | 24 | 25 | 26 |
| 27 | 28 | 29 | 30 | 31 | | |

**1961 SEPTEMBER 1961**

| S | M | T | W | T | F | S |
|---|---|---|---|---|---|---|
| | | | | | 1 | 2 |
| 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| 17 | 18 | (19) | (20) | (21) | 22 | 23 |
| 24 | 25 | 26 | 27 | 28 | 29 | 30 |

**1961 OCTOBER 1961**

| S | M | T | W | T | F | S |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| 15 | 16 | (17) | (18) | (19) | 20 | 21 |
| * 22 | 23 | 24 | 25 | 26 | 27 | 28 |
| 29 | 30 | 31 | | | | |

**1961 NOVEMBER 1961**

| S | M | T | W | T | F | S |
|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 |
| 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 12 | 13 | (14) | (15) | (16) | 17 | 18 |
| 19 | 20 | 21 | 22 | 23 | 24 | 25 |
| 26 | 27 | 28 | 29 | 30 | | |

**1961 DECEMBER 1961**

| S | M | T | W | T | F | S |
|---|---|---|---|---|---|---|
| | | | | | 1 | 2 |
| 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| 17 | 18 | (19) | (20) | (21) | 22 | 23 |
| 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| 31 | | | | | | |

**1962 JANUARY 1962**

| S | M | T | W | T | F | S |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| 14 | 15 | (16) | (17) | (18) | 19 | 20 |
| 21 | 22 | 23 | 24 | 25 | 26 | 27 |
| 28 | 29 | 30 | 31 | | | |

(17) Regular World Day (RWD)

(18) RWD with highest priority

21 Day with unusual meteor shower activity

[15] Day of solar eclipse

16 17 18 19... World Meteorological Interval (WMI)

14 15 16 17... Regular World Interval (RWI)

(12 13 14 15...) International Rocket Week (IRW)

Notes : (*) WMI in January and October considered most important

(* *) Feb. 15, 1961 : RWD with highest priority



5A—(4 pp.)

The RWD with highest priority are for work which can be undertaken for only one day each month. All of the foregoing examples apply if the rate of 3 days per month is not necessary or proves to be too heavy. A specific example is the recommended program for 1959 and onwards of exchange of copies of original ionograms in ionospheric vertical sounding work, made by the URSI–AGI Committee; this recommends that ionograms for the highest priority RWD each month (and also one disturbed period each year) be sent to World Data Centres for interchange.

3. *Regular World Intervals (RWI)*

These are 10 consecutive days in each quarter year, selected to include the three RWD of the month and also the times of the equinox or solstice. If possible they include days of solar eclipses and meteor showers. The selections have also tried to avoid weekend days and in addition holidays which are widely observed.

The RWI are intended for experiments which for practical reasons cannot be carried on continuously, but for which statistics of seasonal variations are especially needed. Ionospheric drift and high atmosphere wind measurements are two examples. Schedules for interchange of sample detailed data in several disciplines have made use of the RWI. In some network observational programs, both the RWI and RWD are used, in order to get the variations throughout the year but with improved statistics at the equinoxes and solstices.

4. *World Meteorological Intervals (WMI)*

The WMI are 10 consecutive days each quarter year, but displaced one month from the equinoxes and solstices. They are intended to cover the times of marked seasonal change in certain meteorological phenomena which tend to come about a month after the equinoxes and solstices. They have been chosen through WMO and COSPAR channels as 16–25 January, April, July and October, with January and October designated by COSPAR as the more important ones.

The WMI now are primarily periods for carrying out synoptic meteorological rocket programs, with stations obtaining atmospheric profiles up to 50 km or more, at least once daily during the 10-day intervals. The WMI have also been used during and since the IGY for balloon sounding programs either with special instruments or launchings to unusually high balloon altitudes.

5. *International Rocket Weeks (IRW)*

These have been chosen for 1961 by COSPAR as (I) 12–18 February and (II) 16–22 July. IRW-I includes the time of the total solar eclipse of 15 February and is intended for study of solar effects. IRW-II was selected for study of (northern hemisphere) summer atmospheric structure.

The IRW provide two periods during the year when rocket studies of the atmosphere and of the sun will be on as nearly a synoptic basis as is possible at this stage of the science and technology. More detailed recommendations by Working Group 2 of COSPAR, J. Bartels, Chairman, appear in the *COSPAR Information Bulletin*.

6. *Other special days*

The 1961 Calendar marks the days of solar eclipses—15 February (total) and 11 August (Annular). Some special programs may be expected to be carried out in appropriate parts of the world to study eclipse effects on the earth's atmosphere. It is especially important that the record of solar activity during and near the times of eclipses be as full as possible. Many solar activity observatories issue specially detailed reports of their observations on



VOL.
20
1961

eclipse days to assist the interpretation of the geophysical efforts. Ionospheric stations customarily increase their observing programs on eclipse days even if the magnitude of eclipse at their location is small.

Also shown on the 1961 Calendar are days when meteor shower activity is unusually high. Geophysicists using meteor techniques often enhance their observing programs on these days. Attention is also called to these days in case ionization produced by meteors may account for unusual effects in other geophysical experiments.

7. *Special Intervals not appearing on* 1961 *Calendar*

The International Geophysical Calendar marks only those dates which can be selected long in advance, either by general agreement or, in the case of eclipses and meteor showers, by reliable long-term prediction. Periods of great magnetic, auroral and ionospheric disturbance are also of great geophysical interest and world-wide coordination of observation is clearly desirable. This is also done under the auspices of the International World Day Service, in close collaboration with the URSI Central Committee on URSIgrams, in the program for the immediate designation of Geophysical Alerts and for selection on a current basis of Special World Intervals (SWI). Arrangements for receipt of such information by telegram or radio broadcast can be made, as may be practical, with one of the solar-geophysical Regional Warning Centres, whose telegraphic addresses are as follows: (Western hemisphere) AGIWARN WASHINGTON (U.S.A.); (Western Pacific) AGI KOKUBUNJI (Japan); (Eurasia) NIZMIR MOSCOW (U.S.S.R.); (Western Europe) either IONOSPHARE DARMSTADT (G.F.R.) or GENTELABO PARIS (France) or AGI NEDERHORSTDENBERG (Netherlands). The meteorological telecommunications network coordinated by WMO carries such information once daily soon after 1600 U.T. Description of the GEOALERT and SWI plan can be obtained from these centres or from the IWDS Secretary. Many geophysical stations increase their programs or carry on special experiments during disturbed periods; the GEOALERT and SWI program serves to coordinate this on a world-wide basis and is especially useful for stations not near the auroral zones where the beginning of a major disturbance may not be immediately apparent from local observations.

The IWDS, in close collaboration with the URSI Central Committee on URSIgrams, also fosters arrangements for prompt notification of major solar flare events which have important and sometimes long lasting geophysical effects. These notifications are also done through the Regional Warning Centres.

8. *Calendar Records*

A summary record of significant solar and geophysical events is being prepared as a Calendar Record for the IGY period and also for 1959. This work is now under IWDS auspices. If these volumes prove to serve a useful purpose, similar Calendar Records may be compiled for 1960 and 1961.

9. The International World Day Service (IWDS) was established in 1958 by the International Council of Scientific Unions (ICSU) and is administered by the International Scientific Radio Union (URSI), 7 Place Emile Danco, Brussels 18, Belgium. The IWDS Steering Committee consists of A. H. Shapley (URSI), M. Nicolet (International Union of Geodesy and Geophysics—IUGG), and J. F. Denisse (International Astronomical Union—IAU) with R. Coutrez (URSI) as Secretary. The IWDS obtains nominal support from the ICSU Federation of Astronomical and Geophysical Services (FAGS). Mr. Shapley serves as spokesman for IWDS and its correspondent to other ICSU groups such as the International Committee on Geophysics (CIG) and the Committee on Space Research (COSPAR).

The International Geophysical Calendar for 1961 has been drawn up by A. H. SHAPLEY and J. V. LINCOLN in consultation with URSI, IUGG and IAU, both directly and through the CIG and COSPAR. Recommendations also have come from representatives of the WMO and from interested individual scientists. A similar Calendar was issued for 1960 along the lines of the calendars for the IGY and IGC 1959 issued under the auspices of the ICSU Special Committee for the International Geophysical Year (CSAGI), and described in the IGY Instruction Manual for World Days and Communications, *IGY Annals* Vol. VII, Pergamon Press, 1959.

*Chairman, IWDS Steering Committee* A. H. SHAPLEY
*Boulder, Colorado, U.S.A.*

---

## A note on polar black outs

(*Received* 24 *September* 1960)

ONE OF THE CHARACTERISTICS of a polar absorption event which has been the source of a considerable amount of discussion is the night-time recovery phenomenon. Absorption of cosmic radio noise begins to decrease as soon as the earth's shadow line rises above the ground and reaches its normal value when the shadow is about 55 km high. This recovery is attributed to the formation of negative ions which readily lose their electrons in the sunlight (BAILEY, 1959; HULTQVIST and ORTNER, 1959; ERICKSEN *et al*, 1960; REID and LEINBACH, 1960). A difficulty is connected with the fact that an increase in detachment continues until ground sunrise has occurred. This is interpreted by some authors as meaning that ionization from the flux of solar protons must be occurring at very low altitude well below 50 km (HULTQUIST and ORTNER, 1959). Objection to this interpretation is based on the fact that physically it is difficult to believe that the protons in great numbers penetrate almost to sea level, and to the fact that below 50 km the electron collisional frequency becomes so high that an extraordinarily large electron density would be needed at such levels to produce the observed absorption (ERICKSEN *et al.*, 1960; REID and LEINBACH, 1960). Thus it is felt that a screening layer at high altitude must exist for the detaching radiation. The favourite candidate for this screening is the ozone layer which would imply that the effective radiation is in the ultra-violet. In turn this means that the negative ion involved would not be $O_2^-$ (or $O^-$), from which the electron is easily detached in visible light, but some species such as $NO_2^-$ (whose affinity—1·6 eV—is still too low) or $O_3^-$ (REID and LEINBACH, 1960). A difficulty with $O_3^-$ is to find enough $O_3$ above 50 km, especially since $O_3^-$ has never been observed in nature.

The purpose of the present note is to point out that because of scattering and refraction in the troposphere and because of absorption by ozone (Chappuis bands) appreciable attenuation of visible light occurs during twilight at altitudes far above the surface of the earth. Thus the effective shadows line will lie well above the geometrical one. In fact transmission functions for yellow light in the atmosphere have been computed by several authors in connexion with the problem of the sodium twilight flash* (HUNTEN, 1954; BLAMONT

* A controversy exactly the analogue of the present once existed over whether the earth's shadow or that of the ozone screening layer was the one involved in the case of the twilight flash. In fact the present treatment is an exact analogue of that applied by HUNTEN (1954) to the sodium twilight flash.

*et al.*, 1958). A group of these is plotted for example in the paper by BLAMONT *et al.* (1958). The transmission begins to decrease for solar rays with an impact parameter of about 50 km and is down to 50 per cent for an impact parameter of about 25 km. Thus when the solid earth shadow is 14 km high (5° angle of solar depression) the intensity of the yellow sunlight illuminating the atmosphere 50 km above the earth has only half its maximum value. A family of transmission functions giving the relative intensity of the yellow light arriving at a point in the atmosphere as a function of its altitude for various angles of solar depression is plotted in Fig. 1. The percentage of solar noise absorption as measured by ERICKSEN *et al.*

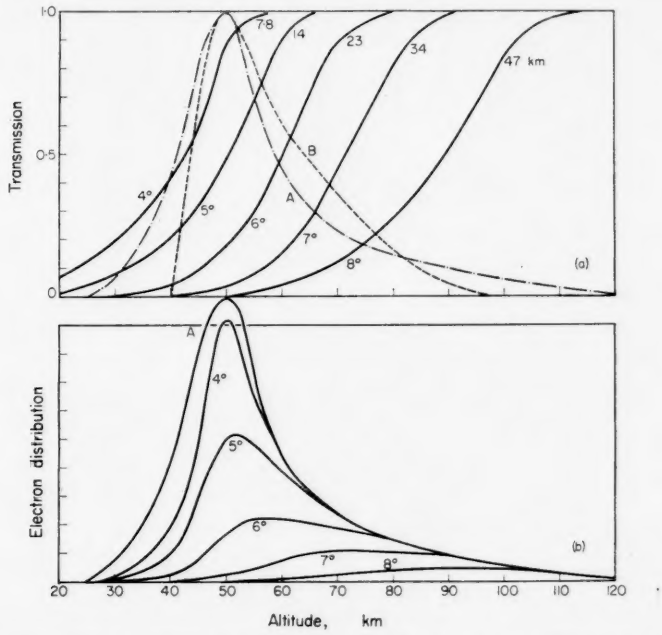


Fig. 1. (a) Transmission function of atmosphere for yellow light as function of altitude for various angles of solar depression. The height of the earth's shadow line is indicated also for each function. Curves A and B give the distribution in arbitrary units of ionization produced by cosmic ray particles. (b) Product of the transmission functions and curve A from (a). These curves give the distribution of free electrons for different distances of the sun below the horizon.

(1960) is plotted as a function of the angle of solar depression in Fig. 2. The general trend of these points is quite similar to the transmission function at 50 km except for having a much longer tail. This suggests that if the free electron density in twilight be taken as proportional to the product of the transmission function and the density of ionization produced by the incident solar particle flux, the peak of the ionization must be close to 50 km. A curve (A) in Fig. 1 gives a proposed distribution of ionization as function of altitude. For a given angle of solar depression this function is multiplied by the appropriate transmission function to yield the distribution of free electrons shown in the bottom part of the figure. Each of these distribution functions is integrated and the ratio of the integral to the area under the curve A taken as proportional to the percentage of maximum radio noise absorption. The result is plotted in Fig. 2. For comparison the result for a somewhat different distribution B also peaked at 50 km but with less pronounced tails toward high and low altitude is also



VOL. 20 1961

shown. It is easy to verify that the peak in the distribution cannot be shifted far from 50 km without causing a serious disagreement with the experimental results.

It is probably not a coincidence that the peak in the ionization affecting the absorption is found at 50 km while the ability of free electrons to produce absorption drops sharply below 50 km. The curve A is really the distribution of efficacious ionization containing the instrumental response curve which decreases rapidly below 50 km.

The above considerations are rather crude and are meant only to demonstrate roughly that the twilight effects are probably explicable in terms of a reasonable distribution of negative ions and that these need be nothing other than ions of oxygen, since the visible spectrum is adequately screened by the troposphere and stratosphere to be the detaching agent. An exact calculation would have to take into account the rates of ionization, of

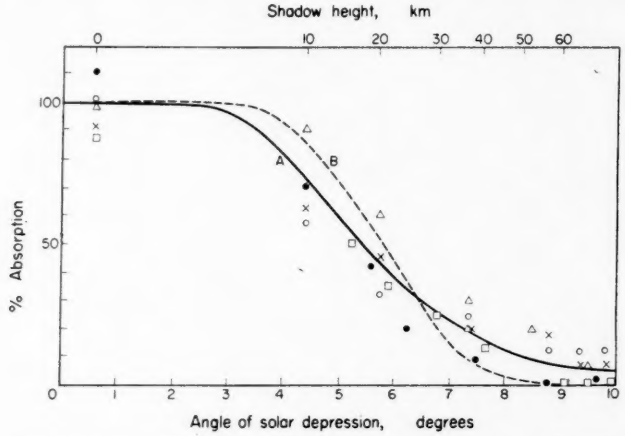


Fig. 2. The points give the galactic noise absorption observed at different Norwegian stations (three) for different angles of solar depression. Curves A and B are the predicted absorption corresponding to cases A and B of Fig. 1(a). The abscissa also is marked in terms of the height of the earth's shadow.

attachment and detachment in a self consistent way. In this connexion, however, I would like to draw the attention of ionospheric physicists to the recent discovery that negative ions of the atmospheric gases may be formed by direct transfer of an electron from a fast incoming hydrogen atom with a very large cross-section (DONAHUE and HUSHFAR, 1959, 1960; CURRAN and DONAHUE, 1960). The process would lead to the formation of $O^-$ by the dissociation reaction:

$$H + O_2 \rightarrow p + O^- + O$$

as well as (especially at lower energies) $O_2^-$ in

$$H + O_2 \rightarrow p + O_2^-.$$

Although the cross-sections for these have not yet been measured, experience with CO indicates that the cross-section will be as large as $2 \times 10^{-16}$ cm$^2$ for hydrogen atoms between 0 and 50,000 eV (or higher). Since the incoming beam in the important part of its range is composed of protons and large numbers of hydrogen atoms formed by charge transfer, it becomes necessary to decide, by comparing the direct negative ion formation rate with that due to attachment of free electrons after ionization, what is the relative importance of the two processes.



VOL. 20 1961

However it is to be noted that the direct transfer process leads directly and predominantly perhaps to $O^-$ formation instead of $O_2^-$.

Thus it would appear that the twilight effects in polar black outs are explicable in terms of the detachment of electrons from negative oxygen ions in the region of 50 km and higher by visible light. The ions involved may well be $O^-$ as well as $O_2^-$, however, because of the importance of direct formation of negative ions in oxygen by the incoming beam particles in their neutral state.

*Acknowledgements*—The research reported here was supported in part by the U.S. Army, Office of Ordnance Research and by the Office of Naval Research. The author wishes to thank the John Simon Guggenheim Foundation for making possible the sojourn at the Service de'Aéronomie, Observatoire de Meudon during which this note was written.

T. M. Donahue

*The University of Pittsburgh*
*Pittsburgh, Pennsylvania*

References


| | | |
|---|---|---|
| Bailey D. K. | 1959 | *Proc. IRE* **47,** 255. |
| Blamont J. E., Donahue T. M. and Stull V. R. | 1958 | *Ann. Géophys.* **14,** 253. |
| Curran R. and Donahue T. M. | 1960 | *Phys. Rev.* **118,** 1233. |
| Donahue T. M. and Hushfar F. | 1959 | *Phys. Rev. Letters* **3,** 470. |
| Donahue T. M. and Hushfar F. | 1960 | *Nature, Lond.* **186,** 1038. |
| Ericksen K. W., Holt I. and Landmark B. | 1960 | *J. Atmosph. Terr. Phys.* **18,** 78. |
| Hultqvist B. and Ortner J. | 1959 | *Planet. Space Sci.* **1,** 193. |
| Hunten D. M. | 1954 | *J. Atmosph. Terr. Phys.* **5,** 44. |
| Reid G. C. and Leinbach H. | 1960 | University of Alaska Report. |


---

## Electrode effect measurements above the sea

(*Received* 14 *October* 1960)

Until now atmospheric electric investigations have not proved the existence of the electrode effect over land. Theoretically, in the condenser Ionosphere–Earth there will be produced an electric space charge at the electrodes which causes an increase of the atmospheric electric potential gradient at the ground. In a condenser containing ions, the potential gradient $E_0$ at the electrode must be double $E_\infty$ at a greater distance, because the transport of the current to the electrode is accomplished by current carriers of only one sign. That means

$$E_0 \cdot \lambda_+ = E_\infty(\lambda_+ + \lambda_-).$$

The electrical conductivity of the air, produced by positive small ions, $\lambda_+$, is nearly the same as $\lambda_-$, which is produced by negative small ions. Therefore

$$E_\infty \approx \tfrac{1}{2}E_0.$$

This has never been measured over land.

In order to explain this phenomenon over land Chalmers (1946) assumed that radioactive sedimentary substances on the ground cause a higher conductivity, i.e. $\lambda_{+_0} > \lambda_{+_\infty}$;

for example $\lambda_{+_0} \approx 2\lambda_{+\infty}$. In this case $E_0$ might become $E_\infty$, which would confirm the observations discussed.

This idea led us to investigate the electrode effect above the sea. Here one can be sure of finding radioactive substances in the air, but not at the electrode (the surface of the water).

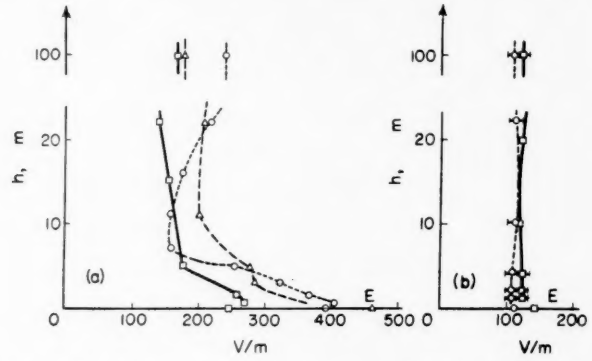


Fig. 1. Electrode effect over sea and land. ($E$ = potential gradient as a function of the altitude $h$ at different times).
(a) Lake Constance, 5 April 1960, cloud 2/10 Cs SW-1, visibility 30 km, $t$ air 17·5°C, $t$ water 8°C, r.h. 39 per cent.
(b) Plane meadow, 6 April 1960, cloud 1/10 Cu SW-2, visibility 50 km, $t$ air 19·5°C, r.h. 51 per cent.

Figs. 1 and 2 show the results of measurements which have been made on Lake Constance (60 km long, 14 km broad) with the atmospheric electric radiosonde of MÜHLEISEN and FISCHER (1958), compared with similar measurements over land. These prove the presumed

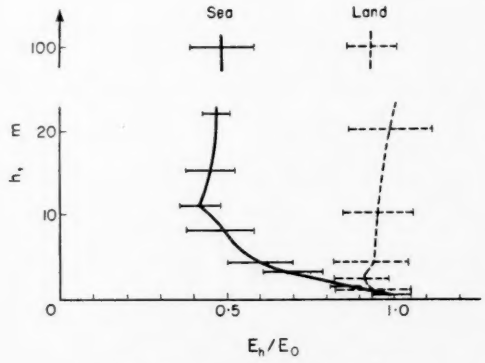


Fig. 2. Mean values of $E_h/E_0$
($E_0$ = potential gradient at the surface;
$E_h$ = potential gradient at altitude $h$).

existence of the electrode effect over the sea and its extension to an altitude of about 10 m.

Simultaneous measurements of the air–earth current, the concentration of positive and negative small ions, and the space-charge density gave good agreement. These results will be given later after having been completed by measurements in different weather situations.



VOL.
20
1961

These results pose several problems:
(1) To what altitude does a potential gradient correspond, having been measured from a ship (Carnegie, etc.)? And to what extent are the reduction factors on level ground to be regarded as sufficiently exact?
(2) What conditions are found over snow-covered land? (Jungfraujoch, Zugspitze, etc.)

The above results, give Chalmers' hypothesis a greater importance, unless it could not be proved. To investigate the problems, further measurements over land, freshly snowed areas and sea are to be made under different weather conditions.

R. Mühleisen

*Astronomisches Institut der Universität Tübingen*
*Weissenau, Germany*

References

Chalmers J. A. 1946 *Quart. J.R. Met. Soc.* **72**, 199.
Mühleisen R. and Fischer H. J. 1958 *Arch. Tech. Messen* **312**, 229.

---

## Geomagnetic disturbance effects in equatorial $E_s$

(*Received* 18 *October* 1960)

It has been known for some time that at equatorial stations where the quiet day solar magnetic variation $S_q$ in horizontal force is large, day time sporadic-$E$ has some remarkable characteristics. It is very regular in its diurnal variation with the critical frequency following the cos $\chi$ law. It is also abnormally intense and often masks out the regular $E$-layer in routine $h'$-$f$ soundings. The lunar influence on this type of $E_s$ has been discussed by Matsushita (1957). Matsushita has shown that at Huancayo the layer often disappears suddenly for varying intervals during day-time when these intervals are between 0000 and 0300 hours or between 1200 and 1500 hours lunar time. He has also found that this disappearance was earlier than usual in the vicinity of new and full moon, particularly during winter, and that the phenomenon occurred during the first half of the periods when the lunar electric current flow in the equatorial regions was westward. Knecht (1959) found, also at Huancayo, a lunar dependence of first appearance of $E_s$; during equinoxes and December solstice, $E_s$ appeared earlier during new and full moon. Skinner and Wright (1957), from Ibadan data, observed that abnormally low values of $fE_s$ occurred simultaneously with reduced horizontal and vertical force.

At Kodaikanal (magnetic latitude; 1·75°N; geomagnetic latitude, 0·6°N) in addition to the lunar influence, a rather remarkable association of magnetic disturbances with sporadic-$E$ has come to our notice. We find that during the main phases of most magnetic storms, when the horizontal force was considerably subnormal, $E_s$ disappeared suddenly. While the main features of this phenomenon were similar to those associated with lunar currents, the length of time for which $E_s$ remained absent during storm perturbations were, on the average, considerably in excess of those associated with the lunar effect.

In order to see how the effects could be separated from those associated with the lunar current, half-hourly day time (0730–1700 hours, 75°E time) ionosonde records at Kodaikanal for a 3 year period (June 1957–May 1960) were scrutinized. To start with, all cases of $E_s$ disappearance were grouped according to the age of the moon and the number of occurrences

in each group were plotted against lunar age. It will be seen from the resulting curve (Fig.1, lower curve) that the frequency of occurrence of the phenomenon was, in general, maximum around new and full moon with a certain amount of extraneous irregular variation. Next, when the process was repeated, with data for all disturbed days excluded, the behaviour (Fig. 1, upper curve) was more systematic with regard to the age of moon, resulting in better maxima around new and full moon and minima around first and third quarter. This is an indication that the disappearance of $E_s$ was to a considerable extent associated with geomagnetic storm perturbations at irregular intervals, irrespective of the lunar age.

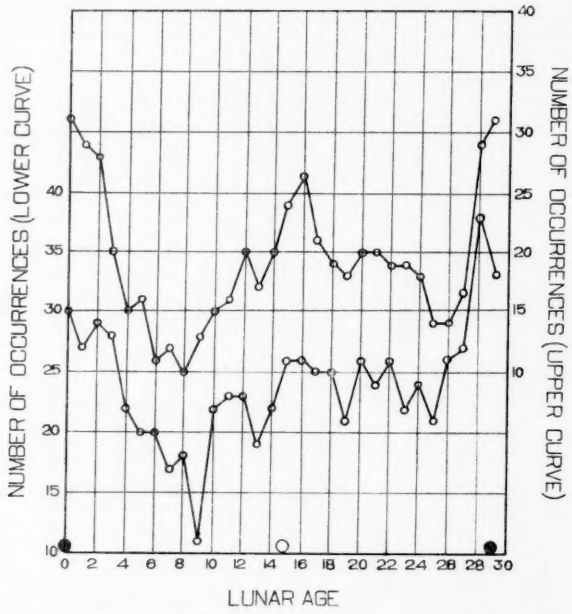


Fig. 1. Variation of number of sudden disappearance of $E_s$ with lunar phase, for all days (lower curve), for days after excluding magnetically disturbed periods (upper curve).

According to MATSUSHITA (1951), the occurrence of intense $E_s$ in equatorial regions has a connexion with the large diurnal range in $H$ at these stations. These ranges are restricted to a narrow belt centred on "zero" dip equator and are explained by the occurrence of an intense eastward current, the so called electrojet, superposed on the normal current responsible for the daily $S_q$ magnetic variation. The $E_s$ disappearance at Huancayo is, according to MATSUSHITA, caused by the westward lunar current which opposes the eastward electrojet. This is confirmed from observations of solar times of $E_s$ disappearance at Kodaikanal where the frequency of disappearance is found to be minimum around 1100 hours which is the time at which the normal $S_q$ and electrojet effects are maximum as seen from Kodaikanal magnetic variations.

In order to obtain preliminary results on the effects of magnetic storms on $E_s$, a superposed epoch analysis was carried out. From inspection of ionograms at $\frac{1}{2}$ hr intervals the time of first disappearance of $E_s$ was taken as zero epoch and the deviations $\Delta fE_s$ were obtained for an 8 hr period around zero epoch. This process was carried out for twenty-seven storms recorded during the 3 year period. For the same period, departures in hori-



VOL. 20 1961

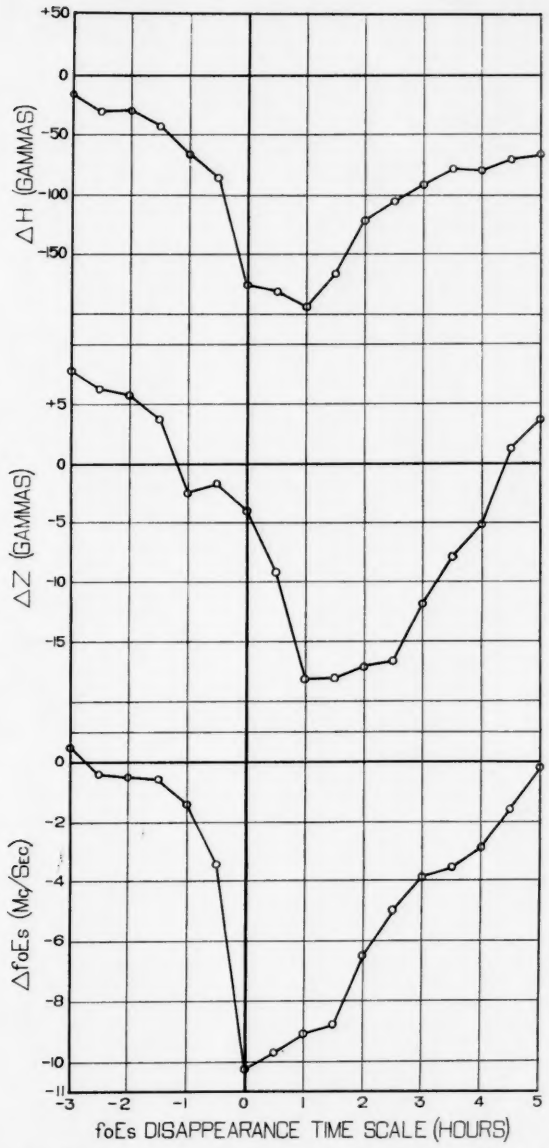


Fig. 2. Departures in $H$ and $Z$ components of earth's magnetic field with those in $fE_s$.

zontal force and vertical force were also obtained and these together with departures in $fE_s$ are plotted in Fig. 2. It will be seen that there is considerable correspondence between $E_s$ disappearance and subnormal values of $H$ and $Z$. The horizontal force drops by over 175 $\gamma$ and vertical force by about 18 $\gamma$ on the average, simultaneously with $E_s$ disappearance.

It is known that the large "effective" conductivity in the dynamo region during daylight (due to inhibition of the Hall currents) associated with abnormally large solar daily

magnetic variation $S_q$ and electrojet is responsible for the formation of equatorial type $E_s$. While the westward current responsible for lunar daily magnetic variation weakens the eastward currents resulting in $E_s$ disappearance, the ring current responsible for $D_{st}$ variation, which is also westward, appears to prevent favourable conditions for the inhibition of the Hall currents at equatorial stations. It appears that this causes a considerable drop in the "effective" conductivity in the dynamo region and results in weakening of the $S_q$ and electrojet currents; consequently we find subnormal values of $H$ and $Z$ and the observed disappearance of $E_s$ for several hours during the main phase of a magnetic storm.

A detailed study covering a period of over half a solar cycle is in progress and the results will be published in due course.

B. N. Bhargava
R. V. Subrahmanyan

*Astrophysical Observatory*
*Kodaikanal, India*

References

Knecht R. W. 1959 *J. Atmosph. Terr. Phys.* **14,** 348.
Matsushita S. 1951 *J. Geomagn. Geoelect., Kyoto* **3,** 44.
Matsushita S. 1957 *J. Atmosph. Terr. Phys.* **10,** 163.
Skinner N. J. and Wright R. W. 1957 *Proc. Phys. Soc. Lond.* B **70,** 833.

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 85 to 93. Pergamon Press Ltd. Printed in Northern Ireland

# Seasonal variation and interpretation of the OH rotational temperature of the airglow*

L. WALLACE
Yerkes Observatory, University of Chicago, Williams Bay, Wisconsin

(*Received* 21 *July* 1960)

**Abstract**—The cascading to lower vibrational levels, following the formation of the excited OH molecules, is accompanied by a shift of the rotational populations to higher rotational levels if there is no collisional redistribution. This effect would cause the rotational temperature for $v = 5$ to be significantly higher than for $v = 9$ for what appear to be reasonable assumptions regarding the primary excitation mechanism. Since the observations indicate that the rotational temperature is the same for all $v$, it appears that collisional redistribution must occur.

A series of thirty-seven spectra of the OH airglow emissions were obtained in the interval December 1958–January 1960. The seasonal variation in the rotational temperatures derived from them is probably less than 30°K and is in phase with the larger variations observed at higher latitudes. The form of the seasonal variation is similar to the form of the seasonal variation of the sodium abundance but not to that of the auroral frequency.

## 1. INTRODUCTION

MEINEL's (1950b) original work on OH gave direct spectroscopic evidence that consistent rotational temperatures could be obtained from the excited level populations but not from the terminal level populations. KVIFTE's (1959b) temperature determinations confirm this result beyond all doubt. Subsequent calculations by MCPHERSON and VALLANCE JONES (1960) indicate that at 90 km the excited molecule would suffer about 200 collisions before radiating. As a result, if the radiation emanates from the 90 km region or lower, as indicated by BATES and NICOLET's (1950) model atmosphere treatment, the rotational temperatures should equal the gas temperatures.

CHAMBERLAIN and OLIVER (1953) noted that the rotational temperature at Thule ($\phi = +76{\cdot}6°$) was distinctly larger than at Yerkes Observatory ($\phi = +42{\cdot}5°$). Numerous later observations [see the summary tabulated by WALLACE (1960a); more recent results are given by KVIFTE (1959a,b) and BLACKWELL *et al.* (1960)] have begun to indicate the form of the variation; a temperature of about 220°K at $\phi = +40$ increasing more and more rapidly to 300° or more at $\phi = +75$, and possibly increasing from $\phi = +40$ toward the equatorial regions. It is also noteworthy that a comparison of the Russian and combined U.S.–Canadian observations indicate that the geomagnetic pole is not the variation pole but that the geographic pole may well be. However, the systematic errors of the various observations and the seasonal variation could seriously affect such an indiscriminate inter-comparison.

The data obtained from rocket studies (JONES *et al.*, 1959) from January to March at an altitude of 75 km indicate that the temperature is about 40° higher at $\phi = +59°$ than at $\phi = +35°$. If the radiation is emitted from an altitude of

* The research reported in this paper was supported by the Geophysics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command, under Contract AF 19(604)-3044 with the University of Chicago.

60–70 km, as indicated by BATES and NICOLET's (1950) theory, it would appear that the latitude variation in the rotational temperatures from $\phi = +35$ to $+60°$ at least might be explained as a change in the gas temperature at a more or less fixed height for the emitting layer. If, on the other hand, the emissions arise from an altitude of around 100 km (KRASSOVSKY, 1957), these rocket observations have no bearing on the problem.

KRASSOVSKY *et al.* (1961) have reported observations pertaining to the seasonal variation in the OH rotational temperatures during the months from October to April. These will be discussed at some length later.

The first part of this paper is an investigation into the likelihood that the OH is in equilibrium at the ambient gas temperature. Although the observation that the rotational temperature is the same for all the vibrational levels tends to indicate that equilibrium has been established, it is not obvious that the lack of variation in this parameter is not the direct result of the excitation mechanism without the intercession of collisional redistribution. Furthermore, both the Einstein coefficient and the altitude of the emitting layer used by MCPHERSON and VALLANCE JONES (1960), although probably accurate enough to warrant their conclusion that equilibrium is established, are only estimates and may be subject to future revision.

The second section is a report on an investigation into the seasonal variation of the rotational temperature.

VOL. 20 1961

## 2. REDISTRIBUTION OF THE ROTATIONAL POPULATIONS

A significant redistribution of the populations of the rotational levels occurs as the natural result of a series of vibrational cascades in the hydrides. This redistribution can be illustrated with a hypothetical excitation mechanism that would populate only the lowest rotational level of the ninth vibrational level of the ground state of OH. As a result of transitions of the type $J'' = J' + 1$, one additional rotational level in each successively lower vibrational level will be populated. As a result of four successive vibrational transitions, the first five rotational levels of the fifth vibrational level will have become populated. Detailed calculations for this example, on the assumption that transitions occur only as $\Delta v = 1$, show that for $v = 5$ the relative populations of the first five rotational levels is reasonably well represented by a temperature of 215°K. It is clear that the temperature increase is greatest when the initial temperature is low, when the rotational constants are large, and when a cascade through several vibrational levels is involved.

From the model calculations to follow it will be seen that this shift of the population to higher rotational levels would cause an observable increase in the rotational temperature as $v$ decreases for any reasonable excitation mechanism for OH. Since this is contrary to the observations it would seem that collisional redistribution does occur.

In these calculations, the usual expression

$$I = (\text{const.})\, \nu^4 i(J) \exp\left[-\frac{b_v hc\, J(J+1)}{kT}\right] \tag{1}$$

was used to determine the relative line intensities within a given sub-band, where $i(J)$ is the intensity factor (or line strength), $b_{v'}$ is the effective rotational constant of the excited level and the other symbols have their usual meaning (HERZBERG, 1950). Although the $^2\Pi$ state involved here is intermediate between HUND's cases (a) and (b), the case (b) factors, which are extremely good approximations to the intermediate case factors (BENEDICT *et al.*, 1953; WALLACE, 1960a), were used in this work. Since, as KVIFTE (1959a,b) has pointed out, the use of the physically meaningful rotational constants, $B_v$, in the above expression is incorrect, the effective rotational constants given by WALLACE (1960a,b) were used.

The calculations were made only for the $F_1$ levels and the rotational temperatures were obtained from the first six rotational lines, $P_1(1)$ to $P_1(6)$, since these are the ones normally used in airglow work.

On the basis of the vibrational intensity distribution MEINEL (1950b) suggested that the OH is formed by some process predominantly in $v = 9$ from which the lower levels are populated by cascade. If it is assumed that $v = 9$ is populated with a Boltzmann distribution at 215° and that transitions occur only as $\Delta v = 1$ (a good first approximation for vibration–rotation bands) without collisional redistribution, the rotational populations are accurately represented by the temperatures given in the column labelled "Model 1" in Table 1.

A subsequent investigation by CHAMBERLAIN and SMITH (1959) indicated that the formation process does not preferentially populate $v = 9$, but rather tends to populate all the vibrational levels equally. The calculations were also performed for a model with this characteristic using the relative transition probabilities given by HEAPS and HERZBERG (1952). For this calculation it was necessary to make some assumption regarding the variation of the rotational temperature with $v$ at which the primary excitation mechanism populates the various levels. (In the work to follow this parameter will be written as $T^*_{\mathrm{rot}}$, in order to differentiate from the rotational temperature, $T_{\mathrm{rot}}$, which describes the sum of the populations due to the primary excitation mechanism *and* the cascading.) The work of BATES and NICOLET (1950) indicates that the excited OH is formed by the ozone-hydrogen reaction whereas KRASSOVSKY *et al.* (1961) favour a combination of the ozone-hydrogen reaction and the $O_2$-hydrogen reaction. It is not known in either case how $T^*_{\mathrm{rot}}$ would vary but it would not be surprising if it were the same for all $v$ or if the rate of populating the rotational levels were the same for all $v$, in which case $T^*_{\mathrm{rot}}(v) \propto b_v$, causing an increase of $T^*_{\mathrm{rot}}$ as $v$ decreases. It was assumed for this model that the former is correct, since with this assumption there is more tendency to counteract the redistribution effect than with the latter. The results of these calculations are given in Table 1 for Model 2.

CHAMBERLAIN and SMITH's (1959) results are of a preliminary nature because the relative vibrational transition probabilities used by them (HEAPS and HERZBERG, 1952) may require substantial revision. It would seem unlikely, however, that such a revision would change the trend of the relative vibrational populations from $v = 9$ to 5 by more than a factor of two. If it is assumed that the variation of $N(v)$ is altered in a smooth fashion such that $N(9)/N(5)$ is decreased by a factor of two, as a result of either revisions of the transition probabilities or new observations, the rate of population due to the primary excitation mechanism alone

increases approximately as 1·0:1·4:1·9:2·5:2·8 as $v$ decreases from 9 to 5, where it was assumed that the ratio of the relative transition probabilities $A_{v,v-1}:A_{v,v-2}:A_{v,v-3} = 6:3:1$ and that $\sum_{v''} A_{v,v''}$ is the same as determined by HEAPS and HERZBERG. The results for this third model, in which $T^*_{\text{rot}}$ was taken as a constant, are also given in Table 1.

A fourth model, which was the same as the third except that $T^*_{\text{rot}}$ was taken as 215, 205, 195, 185, 175°K for $v$ running from 9 to 5, was also calculated.

It may be that none of these models is a good representation of the actual excitation mechanism but the purpose of these calculations was to show that only for excitation mechanisms with characteristics differing considerably from those

Table 1. Results of the redistribution of the rotational populations

| $v$ | $T_{\text{rot}}$(°K) | | | |
|---|---|---|---|---|
| | Model 1 | Model 2 | Model 3 | Model 4 |
| 9 | 215 | 215 | 215 | 215 |
| 8 | 266 | 232 | 230 | 223 |
| 7 | 319 | 250 | 244 | 228 |
| 6 | 376 | 270 | 255 | 233 |
| 5 | 436 | 290 | 266 | 242 |

deduced from the observations and from intuitive guesses as to the variation of $T^*_{\text{rot}}$, would $T_{\text{rot}}$ show negligible variation with $v$.

None of the observations indicate an increase of $T_{\text{rot}}$ as $v$ decreases while those of KVIFTE (1959b)—215 and 214·5°K for $v = 9$ and 6 respectively—appear to deny emphatically any such variation. Consequently it is necessary to conclude that unless our ideas about the behaviour of the excitation mechanism are greatly in error, a collisional redistribution must occur to erase this variation. The only exception to this behaviour seems to be the observation by KRASSOVSKY *et al.* (1961) that $T_{\text{rot}}$ decreases as $v$ decreases. This result could be taken to mean that equilibrium is not established and that the excitation mechanism is very different from those considered here, or that the emission extends over a range of altitudes over which the temperature and relative intensities of the bands also vary considerably. The latter alternative would seem to be more reasonable.

## 3. SEASONAL VARIATION OF THE OH ROTATIONAL TEMPERATURE

### 3.1. *Observations*

The observations were obtained with the 9 in. aperture, $f/0{\cdot}8$ Meinel spectrograph. With this equipment a well-exposed OH spectrum can be obtained on Eastman I–N plates, ammonia hypersensitized according to MEINEL's (1950a) recipe, in about 30 hr during the winter. During the summer, however, with the shorter nights and drop in plate sensitivity owing to increased reciprocity failure, it is difficult to obtain one good plate a month with this technique. After fruitless



VOL. 20 1961

attempts to obtain observations during the summers of 1957 and 1958 the ammonia-hypersensitizing technique was set aside and pre-exposure was tried. Using the latter process adequate seasonal coverage was obtained from December 1958 to January 1960. Of the ninety-eight plates taken for this study thirty-seven were of high enough quality to be used.

The extremely useful report by BOWEN and CLARK (1940) indicated that ammonia hypersensitizing was more effective than pre-exposure in increasing the sensitivity of infra-red materials to low light levels. This result is readily verified when the pre-exposure time is of the order of seconds but when the pre-exposure time is less than $10^{-2}$ sec the sensitivity of the plate is increased considerably to a subsequent exposure at low light levels for long times. The most exhaustive reports on this method of reducing the effects of reciprocity failure are those by HAUTOT and SAUVENIER (1951) and MULIARCHIK and PETROVA (1957). In addition, tests performed at Yerkes Observatory indicated that a $10^{-3}$ sec pre-exposed I–N plate was decidedly more sensitive to subsequent exposures of 4 hr duration than an ammonia-hypersensitized I–N plate. The ratio of the sensitivity increase was ill-determined, probably because of lack of finesse with the ammonia process, but appeared to be in excess of a factor of two. Similar tests on Eastman I–M and I–Z plates indicated that the ammonia process gave much better results than pre-exposure.

The I–N plates used for the airglow study were pre-exposed with a single flash of a commercially available stroboscopic lamp. According to the published characteristics of the lamp and the circuitry involved the lamp flash had a duration of $2 \times 10^{-6} - 5 \times 10^{-6}$ sec. The plates were pre-exposed to give a fog density of about 0·3 as recommended by BRANSCOMB (1951). A plate to be used for intensity calibration was pre-exposed and processed simultaneously with each airglow plate. Both plates were treated as much as possible in the same manner.

The intensity calibration plates were, at first, spectra of an incandescent lamp with an optical step wedge placed across the slit of the spectrograph. This technique appeared to provide good data for intensity calibrations but heavy enough exposures indicated the presence of extraneous light adjacent to the spectrum. Subsequent tests showed that this light was probably due to multiple reflections in the correcting plate of the camera and that the effect amounted to an effective increase in the transparency of the most transparent step of the step filter by about 15 per cent. Corrections were made for this effect and in subsequent calibrations an argon-neon lamp was used in place of the incandescent lamp.

The spectra were reduced to an intensity scale in the usual manner. The continuum, apparently a combination of starlight, moonlight, and incandescent and fluorescent lights, was interpolated between the OH lines and subtracted from the OH line intensities. A standard incandescent lamp operating at a known colour temperature was used to correct for the relative spectral sensitivity of the plate and the transmission of the optical system. These corrections amounted to no more than 18 per cent for any one band.

The airglow spectra were exposed with the eastern sky, about 15° above the horizon, imaged on one half of the slit. The other half, on which the northern sky was imaged, was quite frequently contaminated with aurorae.

3.2. *Temperature determinations*

The spectra were identical in dispersion (~60 Å/mm) to that illustrated by CHAMBERLAIN and ROESLER (1955). It was clear that at most only the $P_1$ and $P_2$ branches of the 5,1 and 6,2 bands and the $P_1(2)$, $P_1(3)$, and $P_1(5)$ lines of the 9,4 band were adequately resolved and bright enough for temperature determinations. Of these the $P_1(6)$ and $P_2(6)$ lines of the 5,1 band fell in a region where it was difficult to interpolate the continuum. In addition, the $P_2(4)$ line of the 5,1 band was often abnormally weak. Because of this anomaly, for which no satisfactory explanation was found, and the difficulty in using the $P_2(6)$ line, the $P_2$ branch of

Table 2. Summary of the observations

| Date | 6,2 | | 5,1 | 9,4 | Plate mean |
|---|---|---|---|---|---|
| | $P_1$ | $P_2$ | $P_1$ | $P_1$ | |
| 7.xii.58 | 213 ± 5 | 209 ± 16 | 187 ± 2 | 201 ± 12 | 191 ± 5 |
| 9.xii.58 | 218 ± 16 | 224 ± 13 | 236 ± 21 | 268 ± 16 | 235 ± 11 |
| 10.xii.58 | 221 ± 12 | 246 ± 36 | 250 ± 14 | — | 234 ± 10 |
| 11.xii.58 | 225 ± 4 | 228 ± 15 | 248 ± 3 | 192 ± 18 | 239 ± 7 |
| 12.xii.58 | 223 ± 8 | 213 ± 5 | — | — | 216 ± 5 |
| 13.xii.58 | 225 ± 4 | 233 ± 28 | 273 ± 46 | 250 ± 23 | 226 ± 4 |
| 15/16.xii.58 | 216 ± 4 | 226 ± 28 | 216 ± 10 | 225 ± 29 | 216 ± 3 |
| 6.i.59 | 232 ± 6 | 237 ± 24 | 229 ± 27 | 338 ± 32 | 235 ± 11 |
| 8/9.i.59 | 201 ± 10 | 289 ± 36 | 214 ± 9 | 203 ± 38 | 211 ± 9 |
| 11/12.i.59 | 221 ± 4 | 275 ± 30 | 270 ± 15 | 223 ± 4 | 224 ± 6 |
| 13.i.59 | 225 ± 9 | — | 251 ± 5 | 230 ± 9 | 242 ± 8 |
| 16/17.i.59 | 226 ± 3 | 247 ± 26 | 225 ± 8 | 257 ± 13 | 227 ± 4 |
| 5/6.ii.59 | 236 ± 5 | 252 ± 17 | 242 ± 17 | 292 ± 2 | 239 ± 6 |
| 11/12.iii.59 | 242 ± 12 | 250 ± 10 | 193 ± 12 | 262 ± 17 | 235 ± 15 |
| 13.iii.59 | 233 ± 6 | 276 ± 25 | 263 ± 9 | 228 ± 2 | 230 ± 5 |
| 16.iii.59 | 229 ± 18 | — | 236 ± 16 | — | 233 ± 3 |
| 18.iii.59 | 204 ± 5 | 191 ± 20 | 214 ± 7 | 208 ± 9 | 207 ± 3 |
| 3.iv.59 | 205 ± 4 | 267 ± 31 | 189 ± 9 | 252 ± 34 | 204 ± 6 |
| 4.iv.59 | 205 ± 4 | 241 ± 18 | 203 ± 8 | 216 ± 6 | 209 ± 4 |
| 5.iv.59 | 139 ± 8 | 202 ± 20 | 183 ± 11 | 144 ± 9 | 154 ± 12 |
| 29.iv.59 | 212 ± 3 | 229 ± 35 | 189 ± 19 | 195 ± 1 | 197 ± 1 |
| 8/9.v.59 | 227 ± 5 | 230 ± 27 | 223 ± 5 | 228 ± 15 | 225 ± 1 |
| 1/2.vi.59 | 233 ± 11 | — | 206 ± 3 | 218 ± 1 | 217 ± 3 |
| 3/4.vi.59 | 192 ± 5 | 239 ± 31 | 193 ± 3 | 220 ± 1 | 216 ± 5 |
| 5/6.vi.59 | 214 ± 7 | — | 233 ± 5 | 207 ± 14 | 225 ± 7 |
| 7/8.vi.59 | 213 ± 12 | — | 190 ± 11 | 191 ± 4 | 193 ± 5 |
| 12/13/14.vi.59 | 212 ± 4 | — | 180 ± 12 | 216 ± 14 | 209 ± 7 |
| 1/2.vii.59 | 201 ± 2 | — | 222 ± 13 | 223 ± 16 | 202 ± 3 |
| 30/31.vii.59 | 190 ± 4 | — | 183 ± 5 | 188 ± 34 | 187 ± 2 |
| 1/2.viii.59 | 230 ± 10 | 238 ± 31 | 215 ± 7 | 193 ± 10 | 213 ± 10 |
| 2.ix.59 | 225 ± 3 | 266 ± 30 | 180 ± 12 | 221 ± 4 | 222 ± 5 |
| 29.ix.59 | 211 ± 4 | 266 ± 28 | 170 ± 8 | 211 ± 8 | 205 ± 10 |
| 9/10.x.59. | 216 ± 11 | — | 210 ± 6 | 201 ± 14 | 210 ± 3 |
| 29.x.59 | 220 ± 5 | 241 ± 28 | 237 ± 11 | 215 ± 10 | 222 ± 4 |
| 3.xii.59 | 227 ± 2 | 257 ± 31 | 286 ± 16 | 240 ± 1 | 237 ± 3 |
| 21.xii.59 | 209 ± 8 | 251 ± 32 | 228 ± 31 | 213 ± 13 | 213 ± 5 |
| 3/4.i.60 | 226 ± 4 | 236 ± 13 | 224 ± 10 | 225 ± 36 | 226 ± 2 |



VOL. 20 1961

this band was not used. As a result, only the $P_1(2)$ to $P_1(5)$ lines of the 5,1 band were used. The lines $P_1(2)$ to $P_1(6)$ and $P_2(2)$, $P_2(3)$, $P_2(5)$ and $P_2(6)$ of the 6,2 band were used whenever strong enough. The $P_2(4)$ line was omitted because of the frequent contamination with the $\lambda$8446 line of oxygen.

The temperatures were determined from the slopes of least squares fits to the expression

$$\ln\left[\frac{I}{\nu^4 i(J)}\right] = -\frac{b_v hc\, J(J+1)}{kT} + (\text{const.}) \qquad (2)$$

obtained from equation (1). The resultant temperatures and their mean errors, determined in the usual manner from the residuals of the least squares fits, and

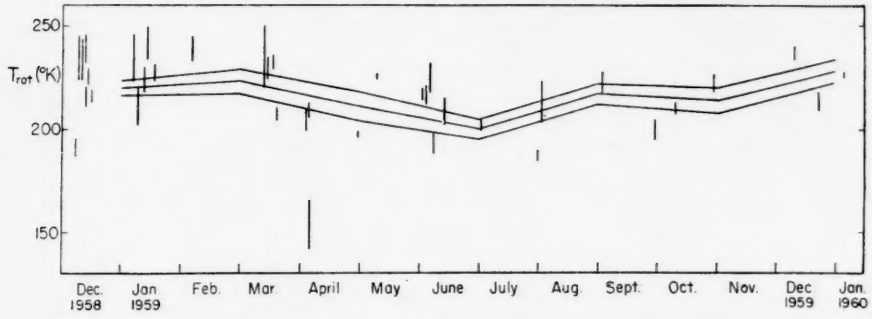


Fig. 1. Rotational temperature variation observed at Yerkes Observatory. The vertical bars indicate schematically the mean error associated with each observation. The bi-monthly means are connected with straight lines, as are their mean errors.

the weighted mean temperatures from each plate are given in Table 2. The weighted mean plate temperatures with their associated mean errors and the weighted mean bi-monthly temperatures and their mean errors are also given in Fig. 1.

### 3.3. *Discussion*

It is apparent from Fig. 1 that if there is a seasonal variation it is probably less than 30°K. The rocket observations of Jones *et al.* (1959) indicate that the seasonal temperature variation at the latitude of Yerkes Observatory at 75 km is probably less than about 20°. Because of the small amplitude of the spectroscopic temperature variation it appears that if the radiation originates from the 75 km region at this latitude there is also small variation in the altitude of the emission.

The bi-monthly mean of the Yerkes ($\phi = 42{\cdot}5°$) temperatures is compared with the mean of the observations of Krassovsky *et al.* (1961) for Yakutsk ($\phi = 62{\cdot}0°$) during the IGY-IGC in Fig. 2. It is clear that the form of the variation in both cases is very similar and that the amplitude of the Yakutsk variation is at least twice as large as that of the Yerkes variation. In the Zvenigorod ($\phi = 51{\cdot}0°$) data, also reported by Krassovsky *et al.*, the scatter, which is larger than in the Yakutsk observations, appears to completely mask any seasonal variation. The fragmentary results of McPherson and Vallance Jones (1960) obtained at Saskatoon ($\phi = 52{\cdot}1°$) indicated that the temperatures of February–March 1957 were 37°K

higher during April–May 1958. The observations from any one station are open to the criticism that the scatter in the observations is comparable to the apparent seasonal variation and that, in some cases, the apparent seasonal variation may be an annual variation, but the general agreement on the form of the variation would appear to be good evidence for the reality of the seasonal variation. The observations also indicate an increase of the amplitude of the seasonal variation with increasing latitude, in the range covered by the observations.

The similarity between the form of the seasonal variation of the sodium abundance (CHAMBERLAIN *et al.*, 1958) and the OH temperature is also evident from Fig. 2.

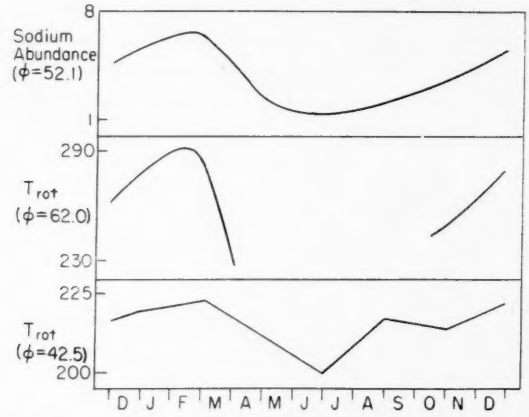


Fig. 2. Intercomparison between the sodium abundance observed at Saskatoon in units of $10^9$ atoms/cm$^2$ column (CHAMBERLAIN *et al.*, 1958) and the seasonal variation of the OH rotational temperature at Yakutsk ($\phi = 62{\cdot}0°$) (KRASSOVSKY *et al.*, 1961) and at Yerkes Observatory ($\phi = 42{\cdot}5°$).

The seasonal variation of the OH temperature clearly does not exhibit the bi-annual periodicity of the auroral frequency variation nor does the February maximum coincide with the Spring maximum in auroral frequency. The suggestion by KRASSOVSKY *et al.* (1961) that the variation is associated with wind systems in the upper atmosphere is attractive but an adequate interpretation will have to wait until these wind systems are better understood.

*Acknowledgements*—It is a pleasure to acknowledge the many interesting discussions with Dr. J. W. CHAMBERLAIN on many aspects of this work. Dr. CHAMBERLAIN also initiated the series of observations used in the work on the seasonal variation. Mrs. VIDYA PESCH and Miss ELAINE SANDBERG obtained many of the spectra.

## REFERENCES

BATES D. R. and NICOLET M. 1950 *J. Geophys. Res.* **55,** 301.

BENEDICT W. S., PLYLER E. K. and HUMPHREYS C. J. 1953 *J. Chem. Phys.* **21,** 398.

BLACKWELL D. E., INGHAM M. F. and RUNDLE H. N. 1960 *Astrophys. J.* **131,** 15.

| | | |
|---|---|---|
| Bowen I. S. and Clark L. T. | 1940 | *J. Opt. Soc. Amer.* **30,** 508. |
| Branscomb L. M. | 1951 | *J. Opt. Soc. Amer.* **41,** 255. |
| Chamberlain J. W., Hunten D. M. and Mack J. E. | 1958 | *J. Atmosph. Terr. Phys.* **12,** 153. |
| Chamberlain J. W. and Oliver N. J. | 1953 | *Phys. Rev.* **90,** 1118. |
| Chamberlain J. W. and Roesler F. L. | 1955 | *Astrophys. J.* **121,** 541. |
| Chamberlain J. W. and Smith C. A. | 1959 | *J. Geophys. Res.* **64,** 611. |
| Hautot A. and Sauvenier H. | 1951 | *Sci. Indust. Photog.* (2), **22,** 201. |
| Heaps H. S. and Herzberg G. | 1952 | *Z. Phys.* **133,** 48. |
| Herzberg G. | 1950 | *Spectra of Diatomic Molecules* (2nd ed.). Van Nostrand, New York. |
| Jones L. M., Peterson J. W., Schaefer E. J. and Schulte H. F. | 1959 | *J. Geophys. Res.* **64,** 2331. |
| Krassovsky V. I. | 1957 | *J. Atmosph. Terr. Phys.* **10,** 49. |
| Krassovsky V. I., Shefov N. N. and Yarin V. I. | 1961 | *J. Atmosph. Terr. Phys.* In press. |
| Kvifte G. | 1959a | *Geofys. Publ.* **20,** No. 12. |
| Kvifte G. | 1959b | *J. Atmosph. Terr. Phys.* **16,** 252. |
| McPherson D. H. and Vallance Jones A. | 1960 | *J. Atmosph. Terr. Phys.* **17,** 302. |
| Meinel A. B. | 1950a | *Astrophys. J.* **111,** 555. |
| Meinel A. B. | 1950b | *Astrophys. J.* **112,** 120. |
| Muliarchik T. M. and Petrova K. I. | 1957 | *Soviet. Astr.* **1,** 108. |
| Wallace L. | 1960a | *J. Geophys. Res.* **65,** 921. |
| Wallace L. | 1960b | *Astrophys. J.* **132,** In press. |


VOL.
20
1961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 94 to 99. Pergamon Press Ltd. Printed in Northern Ireland

# An estimate of the peak sunspot number in 1968

C. M. MINNIS
D.S.I.R. Radio Research Station, Slough*

(*Received* 28 *July* 1960)

**Abstract**—An examination has been made of some statistical properties of the twenty peaks in the relative sunspot number which have occurred since 1750. The frequency distributions of $R$ and $\Delta R$ and the autocorrelation function have made it possible to forecast the magnitude of the next peak which will probably occur in about 1968. It is concluded that there is a probability of at least 0·75 that this peak will lie in the range 110–160.

## 1. The Solar Cycle and Radio Communication

THREE PEAKS in the cycle of solar activity have occurred at intervals of about 10 years since 1928 and each of these peaks, in terms of the relative sunspot number, has been about 40 units higher than its predecessor. The 1958 peak was the highest recorded for at least 210 years and its outstanding features have already been discussed by ELLISON (1957), MINNIS (1958) and NAISMITH (1958).

The solar cycle is responsible for corresponding cyclic changes in the amount of ionization in the ionospheric layers which reflect and absorb the high-frequency radio waves which are used in many of the world's communication systems. The ultimate effect of the ionospheric changes is that a much wider band of frequencies can be used for radio communication when solar activity is high than when it is low and this, in turn, results in greater reliability. On the other hand, exceptionally high solar activity, such as was experienced near the 1948 and 1958 peaks, is often the indirect cause of mutual interference between short-distance radio services which use frequencies between about 40 and 55 Mc/s. This difficulty arises because it is usually assumed, when planning such services, that the radiated signals will not be reflected from the ionosphere and that the maximum range of reception will be limited by the optical horizon.

HITCHCOCK (1959) has pointed out that the development of long-distance high-frequency communications during the past 30 years or so has coincided with the sequence of high-activity solar cycles mentioned above. He has also discussed the possible consequences of a future sequence of low-activity cycles and it appears that the occurrence of such a sequence might lead to radio communication problems the significance of which has not been fully appreciated. In view of this contingency it would be useful, in the forward planning of communication systems, to have an advance estimate of what the smoothed sunspot number is likely to be at its next peak which will probably occur in about 1968. This peak value and the 1958 peak will be referred to later as $R(1968)$ and $R(1958)$.

## 2. The Data on Sunspots

Fairly reliable values of the smoothed sunspot number are available from about 1749, since when the number has passed through twenty peaks including

* Official Communication.

those of 1750 and 1958. A plot of this long series looks deceptively regular but, in spite of many attempts, no satisfactory method of extrapolating sunspot numbers beyond the current cycle has been found. Each individual cycle can be represented approximately by an empirical function, such as that of STEWART and PANOFSKY (1938), but the parameters which determine the amplitude and phase of this function do not seem to follow any easily recognizable sequence which might be used to forecast the characteristics of future cycles.

In these circumstances it seems worth while to investigate the possibility of using some simple statistical characteristics of the sequence of twenty peak values of the smoothed sunspot number from 1750 to 1958 in an attempt to discover at

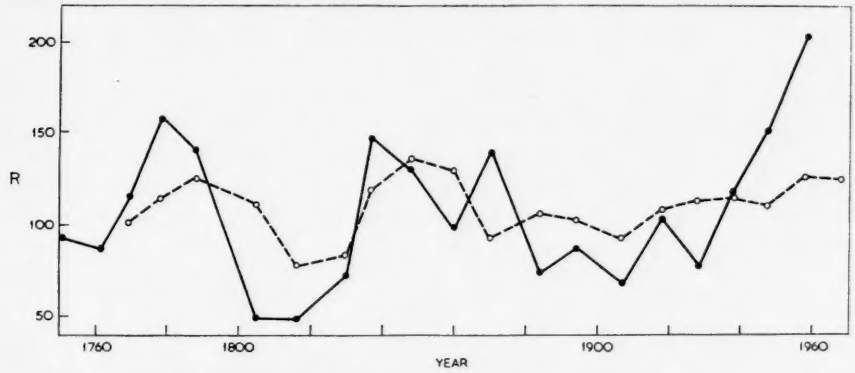


Fig. 1. Smoothed peak values of the relative sunspot number. Full line: observed values. Broken line: forecast values based on the autocorrelation function of $R$ for $r = 1$ and $r = 3$.

least whether it is likely that the 1968 peak will be above the average peak value of 108. This sequence is shown in Fig. 1 and, as it is obviously too short and irregular to allow a precise estimate of its future trend to be made, the objective will be to estimate the probability that the next peak will lie between specified limits. The methods to be used do not depend on cycle-matching procedures. Although this technique has been used successfully by HERRINCK (1959) to forecast the trend of the current cycle, CHADWICK (1959) has shown that it can not be safely used to predict subsequent cycles.

## 3. The Direction Sequence

Before discussing the numerical data in detail, it is important to decide whether there are any grounds for the suggestion that the rising sequence of peak values since 1928 represents the beginning of a trend which may continue for a long time. The differences, $\Delta R$, between the peak values of successive cycles may be either positive or negative in sign. Each difference may be designated $s$ or $d$ depending on whether its sign is the same as or different from that of the preceding difference. Using this terminology, the changes in the peak values in Fig. 1 can be represented by the "direction sequence": *dsdssdsdsddddddddss*. The immediate problem is to estimate the relative probabilities that the next term in this sequence will be $s$, representing an even higher peak in 1968 than in 1958, or $d$ indicating a smaller peak.

The peak values since 1750 vary in a way which suggests that there are upper and lower limits of about 50 and 150 within which they tend to be confined. (Fig. 1). Since both the mean and the most probable values of the magnitude of $\Delta R$ are about 35, the probability of occurrence of a sequence of two or more changes in the same direction, either upwards or downwards, would be expected to be less than that of a sequence of reversals in sign; that is, the probability ($p_1$) of occurrence of $s$ ought to be less than that ($1 - p_1$) of $d$. This hypothesis is supported by the actual frequencies of occurrence of $s$ and $d$ which give $p_1 = 0{\cdot}4$. An obvious extension of this argument leads to the expectation that the probabilities ($p_2$, $p_3$) of occurrence of a second or third $s$, after one or two had already occurred, would be less than 0·4. The observational data are insufficient to allow $p_1$, $p_2$, $p_3$ to be determined accurately, but the best fit between the calculated and observed frequencies of occurrence of $s$ and $d$ is obtained for $p_1 = 0{\cdot}4$, $p_2 = 0{\cdot}3$ or $0{\cdot}35$, $p_3 = 0{\cdot}3$. These values are consistent with the expectation that $p_1 > p_3$.

The expected numbers of occurrences of the four possible pair-combinations of $s$ and $d$ have been calculated using these values of $p$ and are given in Table 1 (Lines (b) and (c)). It must be mentioned, however, that the $\chi^2$ test shows that the fits obtained in these two cases do not differ significantly from the fit which

Table 1. Calculated and observed occurrences of combinations of $s$ and $d$ (1752–1958)

| | | Probability | | | Number of occurrences | | | | $\chi^2$ |
|---|---|---|---|---|---|---|---|---|---|
| | | $p_1$ | $p_2$ | $p_3$ | $dd$ | $ds$ | $sd$ | $ss$ | |
| Calculated | (a) | 0·5 | 0·5 | 0·5 | 4·3 | 4·3 | 4·2 | 4·2 | 2·06 |
| | (b) | 0·4 | 0·35 | 0·3 | 6·3 | 4·2 | 4·3 | 2·1 | 0·19 |
| | (c) | 0·4 | 0·3 | 0·3 | 6·5 | 4·4 | 4·3 | 1·8 | 0·17 |
| Observed | | — | — | — | 6 | 5 | 4 | 2 | — |

results when it is assumed that $s$ and $d$ are equally likely to occur in all circumstances (Line (a)). Nevertheless Lines (b) and (c) are preferred on physical grounds and also because the probability of occurrence of the seven consecutive $d$-terms which appear in the direction-sequence would be less than 0·01 for $p = 0{\cdot}5$.

It is concluded that the best estimates are $p_3 = 0{\cdot}3$, $(1 - p_3) = 0{\cdot}7$ and, since the direction sequence up to the 1958 peak ends in $ss$, it follows that the probabilities that the 1968 peak will be higher and lower than that of 1958 are 0·3 and 0·7 respectively.

## 4. The Frequency Distributions of $\Delta R$ and $R$

The frequency distribution of $\Delta R$ is bimodal and both the mean value and the minimum frequency lie near $\Delta R = 0$. If the signs are disregarded, the mean value of $|\Delta R|$ is 35 and its standard deviation 23. Hence for a negative $\Delta R$ between the 1958 and 1968 peaks, the estimated value of $R(1968)$ is $168 \pm 23$; for positive $\Delta R$ it is $238 \pm 23$. It is also useful to note that $\Delta R$ has exceeded 75 on only one



VOL. 20 1961

occasion out of nineteen since 1750 and that it has never exceeded 92. The subtraction of these values from $R(1958)$ leads to the conclusion that there is a probability of 0·95 that $R(1968) > 128$ and a greater probability that $R(1968) > 111$.

Similar arguments based on the distribution of $R$ lead to the conclusion that the probability that $R(1968) < 159$ is 0·95.

## 5. The Autocorrelation Function

An alternative method of estimating the next peak value of $R$ from those preceding it depends on the use of the autocorrelation function (Fig. 2) which has been calculated for the twenty peak values in Fig. 1. This function represents the linear correlation coefficient for all possible pairs of values of $R$ separated by $r$ cycles ($R_n$, $R_{n+r}$). The form of the function suggests that there may be a period of seven or eight solar cycles in the peak values of $R$ but, except for $r = 1$ and $r = 3$, the correlation coefficients are not significantly different from zero at the 10 per cent level of probability. It is worth remarking, however, that evidence for a period of seven cycles has been obtained by Schove (1955) from the time intervals between successive maxima and minima in solar activity. Since the length of these intervals can be traced back over many centuries using information of various kinds contained in historical records, conclusions based on them may be more reliable than those based on the comparatively short series of peak sunspot numbers.

The regression lines of $R_n$ on $R_{n+r}$ for $r = 1$ and $r = 3$ have been used to forecast $R(1968)$ but, since the correlation coefficients for even these cases are low, the standard errors of estimate (s.e.e.) are fairly large. The two forecasts and their s.e.e. are $154 \pm 38$ for $r = 1$, and $97 \pm 36$ for $r = 3$. The same two regression lines have also been used to forecast each of the peaks from 1761 onwards and a comparison of the forecast and the actual values gives a visual picture of the degree of reliability attainable using this method (Fig. 1).

## 6. The Combination of Individual Estimates

The individual estimates of $R(1968)$ which have been derived in various ways in the foregoing discussions have been collected together in Table 2 together with the probability that each estimate will lie within the quoted limits.

The most objective method of combining these individual estimates is probably to calculate the mean value of Estimates 6–9, weighting each inversely as its variance and, in addition, reducing the weights of Estimates 6 and 7 to 0·3 and 0·7 respectively, in accordance with Estimates 1 and 2, in order that their combined weight should be 1·0, the weight given to both Estimates 8 and 9. The result is that there is a probability of 0·68 that $R(1968)$ will lie in the range 104–218.

The range of the preceding estimate is large because it includes the possibility that $R(1968)$ may be either greater or less than $R(1958)$. The probability that there will be a positive $\Delta R$ between the 1958 and 1968 peaks is 0·3 (Estimate 1), but the occurrence of this event would also imply that $R(1968) > 159$, an event for which the probability is 0·05 (Estimate 3). Hence, assuming the two events to be independent, the probability of the compound event is only 0·02 and, for

practical purposes, it seems reasonable to reject the possibility that $R(1968)$ may exceed $R(1958)$. This assumption implies the rejection of Estimate 6, leaving only Estimates 3, 4, 5, 7, 8 and 9 for further consideration.

The weighted mean of Estimates 7–9 is 149 and the s.d. of the combined

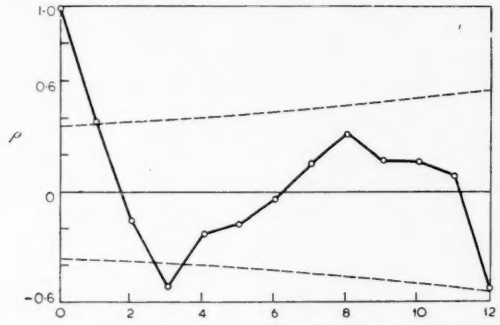


Fig. 2. The autocorrelation function of $R$. The broken lines indicate the 10 per cent significance levels.

estimate is 45. Hence there is a probability of 0·68 that $R(1968)$ will be in the range 104–194. Estimates 3 and 5 together lead to the conclusion that there is a probability of 0·9 that $R(1968)$ will lie in the range 111–159. These two combined estimates are consistent with each other and both suggest that $R(1968)$ will

Table 2. Individual estimates of $R(1968)$

| No. | Method used | $R(1968)$ | Probability |
|---|---|---|---|
| 1 | Direction sequence | $> 203$ | 0·3 |
| 2 | Direction sequence | $< 203$ | 0·7 |
| 3 | Distribution of $R$ | $< 159$ | 0·95 |
| 4 | Distribution of $\Delta R$ | $> 128$ | 0·95 |
| 5 | Distribution of $\Delta R$ | $> 111$ | $> 0·95$ |
| 6 | Distribution of $\Delta R$ | $238 \pm 23$ | 0·68 |
| 7 | Distribution of $\Delta R$ | $168 \pm 23$ | 0·68 |
| 8 | Autocorrelation ($r = 1$) | $154 \pm 38$ | 0·68 |
| 9 | Autocorrelation ($r = 3$) | $97 \pm 36$ | 0·68 |

Table 3. Combined estimates of $R$ (1968)

| No. | Estimates used | $R(1968)$ | Probability |
|---|---|---|---|
| C1 | 3, 5 | 111–159 | 0·9 |
| C2 | 7, 8, 9 | 104–194 | 0·68 |
| C3 | 6, 7, 8, 9 | 104–218 | 0·68 |

probably exceed 108, the mean value of the twenty peaks which occurred between 1750 and 1958.

The combined estimates are summarized in Table 3 and illustrated diagrammatically in Fig. 3(b). The individual estimates in Table 2 are shown in Fig. 3(a).

Since the probabilities and ranges associated with all these estimates depend ultimately on the s.d.'s of samples containing only twenty observations, they may be either greater or less than those which would have been obtained if a much larger volume of data had been available. If the sampling errors are assumed to be of unfavourable sign, the ranges given for Estimates C2 and C3 would have to be increased by 16 per cent and the probability of Estimate C1 would fall to a value which is unlikely to be less than 0·75.

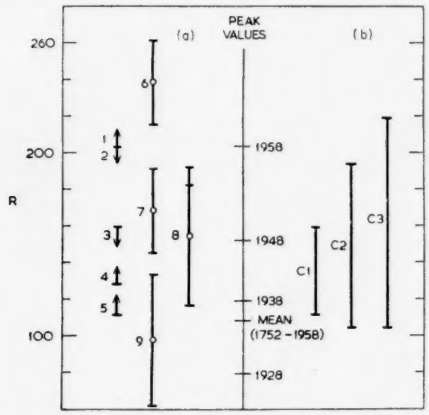


Fig. 3. Forecasts of the smoothed peak sunspot number in about 1968: (a) the individual forecasts (Table 2); (b) the combined forecasts (Table 3).

## 7. Conclusion

There is obviously no justification for making a precise estimate of the smoothed sunspot number at its next peak in about 1968. The range finally adopted is 110–160; this range is based on statistical considerations only and appears to have a probability of at least 0·75 of being correct. Thus conditions in the ionosphere and for radio communications during the solar cycle 1964–1974 would be expected to be comparable with those experienced during the cycles which extended from 1934 to 1944 and from 1944 to 1954.

*Acknowledgements*—The author acknowledges the assistance of Mrs. M. D. Vickers who calculated the autocorrelation function and made the forecasts based on it. The work described was carried out as part of the programme of the Radio Research Board. The paper is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research.

## References


| | | |
|---|---|---|
| Chadwick W. B. | 1959 | *Nature, Lond.* **184,** 1787. |
| Ellison M. A. | 1957 | *Nature, Lond.* **180,** 1173. |
| Herrinck P. | 1959 | *Nature, Lond.* **184,** 53. |
| Hitchcock R. J. | 1959 | *Brit. Commun. Electron.* **6,** 350. |
| Minnis C. M. | 1958 | *Nature, Lond.* **181,** 543. |
| Naismith R. | 1958 | *Nature, Lond.* **181,** 954. |
| Schove D. J. | 1955 | *J. Geophys. Res.* **60,** 127. |
| Stewart J. Q. and Panofsky H. A. A. | 1938 | *Astrophys. J.* **88,** 385. |

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 100 to 109. Pergamon Press Ltd. Printed in Northern Ireland

# Diurnal variation of earth currents at the equator

ROSEMARY HUTTON and R. W. H. WRIGHT
Department of Physics, University College of Ghana

(*Received* 29 *July* 1960)

**Abstract**—Earth current measurements made at Legon, Ghana (5° 38′ N, 0° 11′ W) are analysed for a period of 15 months. From the analysis, the mean diurnal variation is found to be 180 mV/km and the current direction is shown to be at right angles to the nearby coast line. The diurnal variation of earth currents for both quiet days and disturbed days is given and the effect of the different seasons on the quiet day variation is also examined. Good agreement is found between the shape of the diurnal variation curve for earth currents and the diurnal variation of $dZ/dt$, the rate of change of the vertical component of the earth's magnetic field, and the anomalously large value of earth current amplitude is shown on comparison with data for other equatorial stations, to be related to the anomalously large daily range of $Z$ which occurs near the equatorial electrojet.

## INTRODUCTION

CONTINUOUS measurement of earth currents was first made in 1865 at Greenwich, England and since that time investigations have been carried out in Germany, France, Spain, U.S.A., Japan and Russia. There has recently been a marked increase in interest in this method of obtaining information about the earth's magnetic field, especially in connexion with the study of pulsational phenomena. Data reported has, however, been mainly confined to middle latitude stations and for the equatorial region, is limited to the report of work carried out at Huancayo, Peru (GISH and ROONEY, 1930). The diurnal variation of earth currents, at a number of stations, became known about the time of the second International Polar Year 1932–33 and in 1936, GISH (1936) suggested the existence of current whorls to explain these variations. More recent work has, however, shown that the earth current directions are largely dependent on the location of the observatory and in particular, its proximity to the sea and to vast mountain ranges. (BURKHART, 1956; YOSHIMATSU, 1957). The actual magnitude of the diurnal variation is not clearly understood, values reported so far, being of the order of 1–40 mV/km. Again, proximity to the sea is considered to have some effect on these values as shown by observations made in Japan (YOSHIMATSU, 1957).

Despite this dependence of earth currents on the location of the observatory, it is accepted that there is a close relationship between earth current variations and variations of the earth's magnetic field at the same station. BURKHART (1956) has discussed the relationship between earth current variations and the variations of the horizontal component of the earth's field and recently, HESSLER and WESCOTT (1959) have shown that there is a high correlation between earth current activity and the planetary $K$-index of magnetic activity.

The purpose of the present investigation is to provide additional data from an equatorial station and to attempt to relate the results obtained with the earth current results already published from other stations. At the same time, it is known that within a few degrees of latitude from the magnetic equator, anomalously large diurnal variations of the magnetic elements $H$ and $Z$ are



VOL.
20
1961

produced by the equatorial electrojet (EGEDAL, 1948; ONWUMECHILLI, 1959). As Legon is about 400 miles south of the electrojet and within the region where the abnormally large values of $H$ and $Z$ are obtained it is likely that a comparison of the daily variations of the magnetic elements and of earth currents at this station may be of some assistance in the understanding of these phenomena.

## EXPERIMENTAL PROCEDURE

In 1957 an installation for recording earth currents was set up at Legon, Ghana (5° 38′ N, 0° 11′ W) and recordings of earth current potentials have been obtained more or less continuously since that time.

The electrodes were of the porous-pot type, as used by Burkhart and others (BURKHART, 1956). Each consisted of two sheets of copper, placed vertically and at right angles to each other, in a large porous pot which was filled with a mixture of copper sulphate and sawdust. A long piece of iron tubing was fixed over the wooden lid of the pot. This served as a funnel for adding copper sulphate solution from time to time to the pot and as a support for attaching the cable from the electrode to the earth's surface. The electrodes were buried at a depth of 6 ft, where the moisture content of the soil was known to be reasonably constant over long periods of time. Two electrodes were placed along the earth's magnetic meridian, the electrode spacing being 500 m and a second pair of electrodes was situated at right angles to this direction, the electrodes being the same distance apart. They will be referred to as the N–S and the E–W directions. As a result of the various precautions taken, the average electrode potential differences have remained approximately zero since recordings started.

The potential differences between the N–S and E–W electrodes were measured by means of galvanometers whose deflexions were continuously recorded on photographic paper. The recording speed was normally 5 mm/min so that small rapid changes of the earth currents could be detected in addition to the slower diurnal changes which are the subject of this paper.

Calibration was carried out by inserting a known e.m.f. into each line in turn for a short interval of time and noting the deflexion of the traces.

For comparison purposes, earth current measurements have also been made at Tamale, Northern Ghana (9° 25′ N, 0° 51′ W) for a period of 4 weeks and at Ibadan, Nigeria (7° 27′ N, 3° 54′ E) for a similar period of time. In these cases, the electrodes consisted of copper earth rods, 4–8 ft long, and a pen recorder, which operated at a speed of 2 cm/hr was used in place of the fast photographic recording method, employed at Legon. In the cases of Legon and Tamale, the sites chosen for the earth current installations were reasonably level and of uniform geological structure, while at Ibadan, there was a known geological fault in the locality.

## METHOD OF ANALYSIS

Hourly values of the earth currents in the two directions were obtained by visual examination of the records and by estimation of the mean height, during each hour, of each trace above some arbitrary level. From these values, the daily mean was obtained and by subtracting this daily mean from each hourly value, a series of hourly deviations from the daily mean were obtained. In the analysis

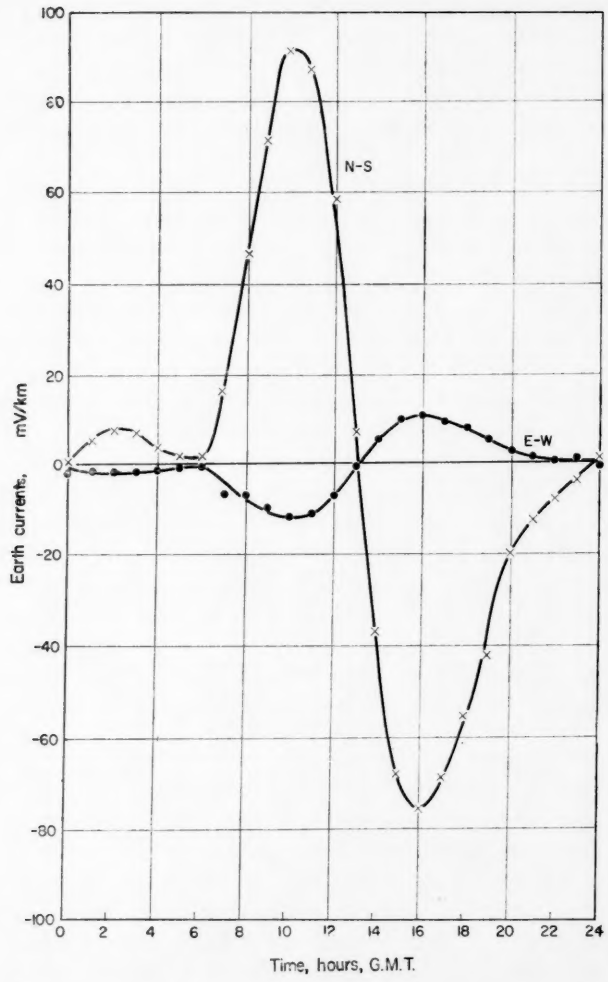


Fig. 1. Diurnal variation of earth current components. Mean of all days, November 1957–October 1958.

only those days have been considered where a complete set of hourly deviations is available.

From this data, the following analyses have been made:

(a) Determination of the mean diurnal variation of earth currents for all days;
(b) Determination of the diurnal variation for quiet and disturbed days;
(c) Determination of the $S_q$-variation for the different seasons;
(d) Comparison of the earth current daily variation with the daily variation of the magnetic elements:
(e) Comparison of the daily earth current amplitude at Legon, Tamale and Ibadan.

*Mean diurnal variation*

The hourly deviations from the daily mean were averaged for each month in the interval November 1957–October 1958, the values being expressed in mV/km. The average of all these hourly deviations is given in Fig. 1, positive values

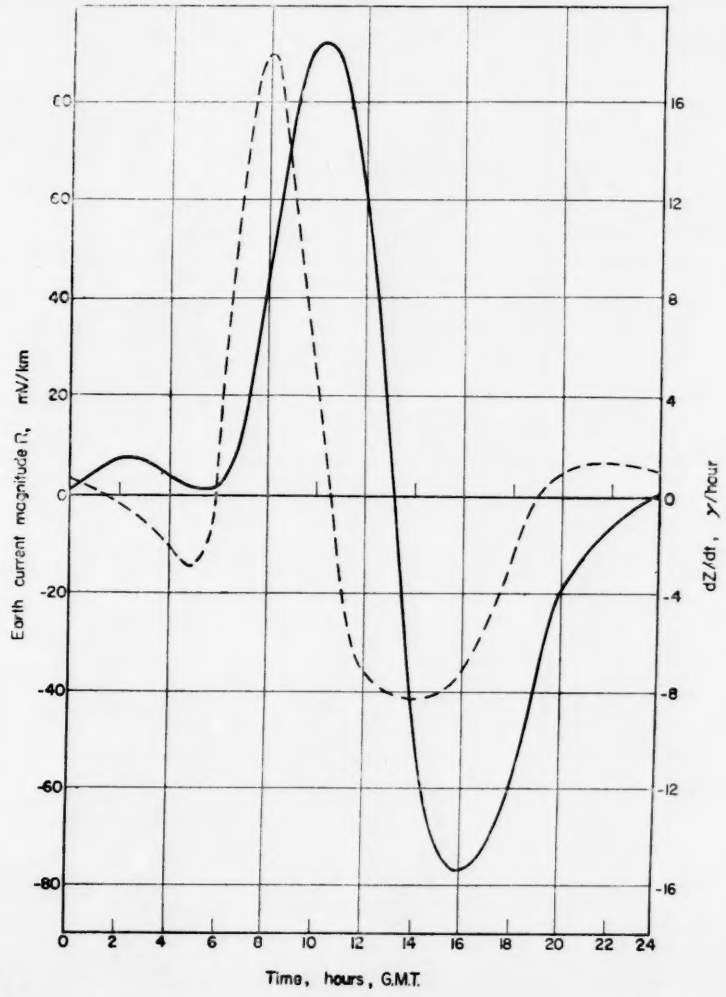


Fig. 2(a). Diurnal variation of (1) earth current magnitude $R$, (2) $dZ/dt$ at Ibadan.

referring to the cases where the N–S trace indicates a flow of current towards the north and the E–W trace indicates a flow of current towards east. The same data has also been compounded vectorially hour by hour giving, (1) a graph of the total magnitude of the potential difference against time of day—Fig. 2(a), and (2) a hodograph, from which the direction as well as the magnitude of the potential difference at any hour can easily be seen—Fig. 2(b). In Fig. 2(a), $R$ has been

considered positive when the resultant current flow is in a northerly direction and negative when in the reverse direction.

The shape of the diurnal variation curve, Fig. 2(a), agrees well with that obtained at Huancayo but its amplitude is approximately 60 times greater (GISH and ROONEY, 1930). It also has a similar shape to the variation of $dZ/dt$ as

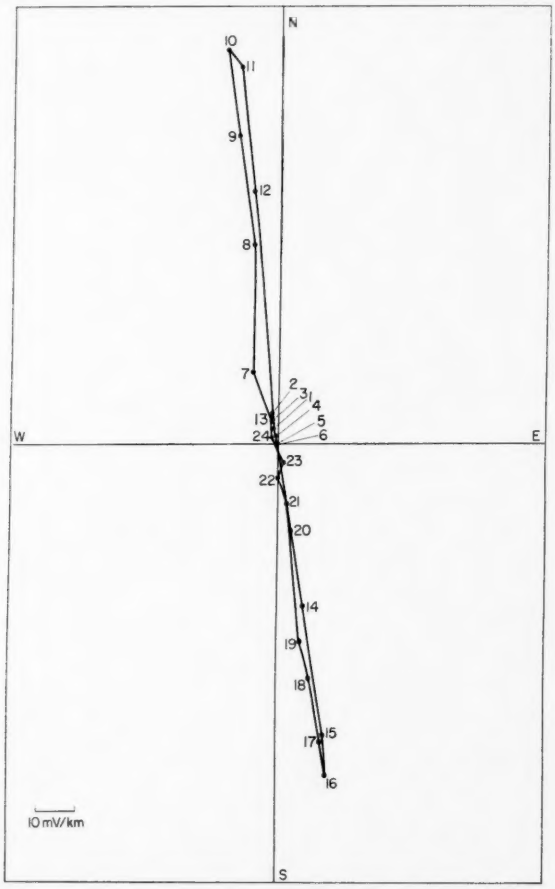


Fig. 2(b). Earth current hodograph.

observed nearby at Ibadan, Nigeria. The variation of $dZ/dt$ at Ibadan is plotted on the same figure. These points will be discussed more fully later.

The hodograph, Fig. 2(b), shows that the current direction is more or less constant and assuming a value of declination $D$ of 9° W of N at Legon, this gives the current direction as approximately 17° west of the geographic north.

Together Figs. 2(a) and (b) show that there is a maximum flow of current northwards in a direction 15° W of north with a value of approximately 90 mV/km at 100 hours and a maximum flow in the reverse direction of about 80 mV/km at 1600 hours. The main current direction is approximately at right angles to the



VOL.
20
1961

coast line which is about 11 km from Legon. This dependence of current direction on an adjacent coast is in good agreement with results published by YOSHIMATSU (1957) for a large number of Japanese stations.

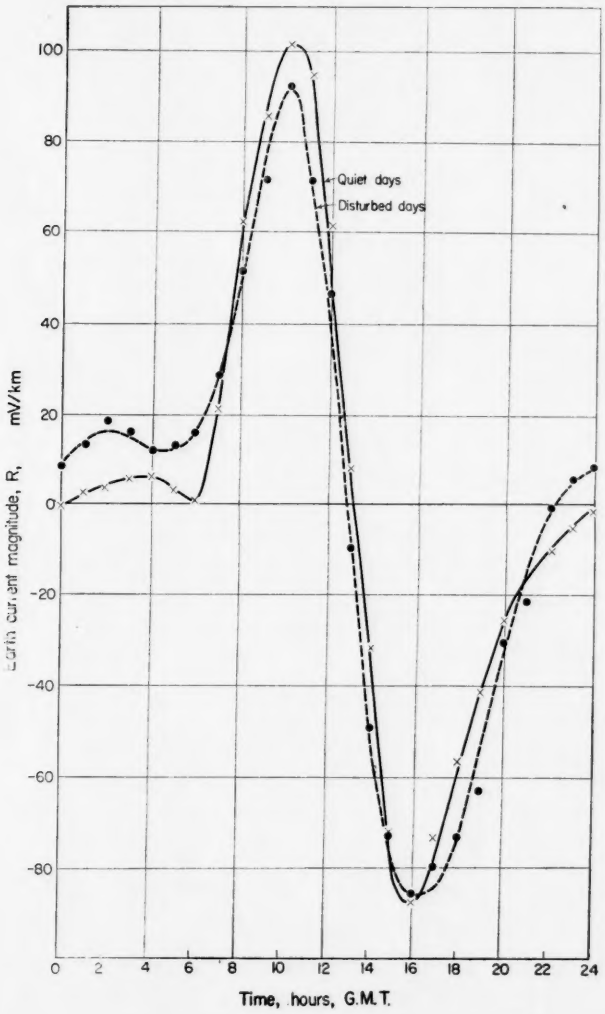


Fig. 3. Diurnal variation of earth current magnitude for quiet and disturbed days.

*Diurnal variation for quiet and disturbed days*

Using the planetary $K$-index as a measure of the activity of the earth's magnetic field, five quiet days have been selected for each month from April 1958 to March 1959 from the days for which a complete set of hourly values was available. These hourly values have been averaged over the whole year to give

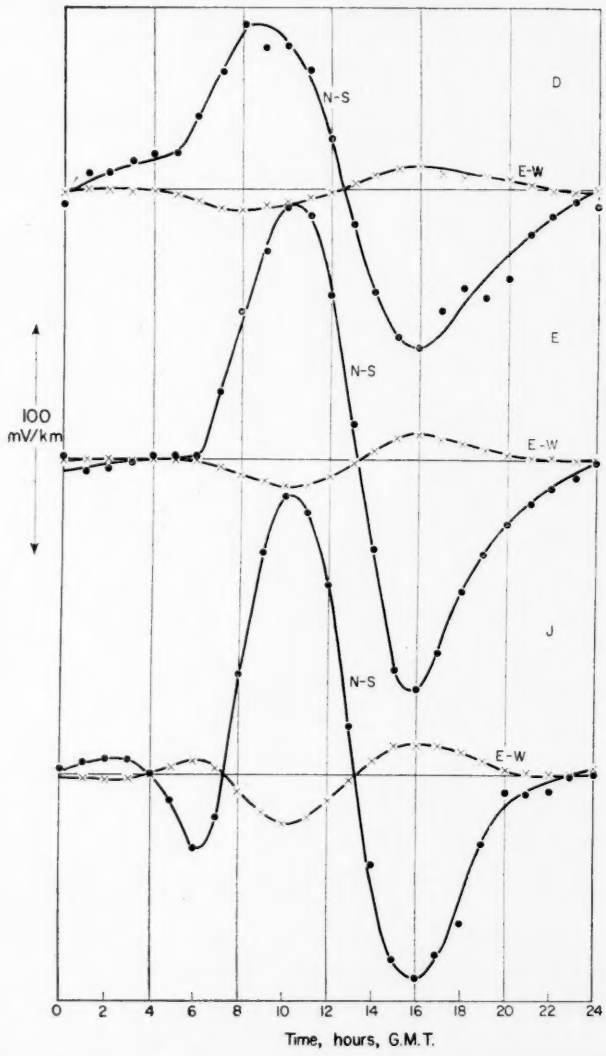


Fig. 4. $S_q$-variation of earth current components for the three seasons.

the mean $S_q$-variation. This is shown in Fig. 4, together with the graph obtained in a similar manner for the five most disturbed days of each month.

The main points of difference between the earth current values on quiet and disturbed days would appear to be:

(1) The daily range on disturbed days is slightly less than on quiet days. It may be pointed out at this stage that the amplitude of $Z$ in the region of the electrojet is less on disturbed days than on quiet days;

(2) The values of the earth currents at night are greater on disturbed days than on quiet days.



VOL.
20
1961

*The $S_q$-variation for different seasons*

The value of the $S_q$-variation of earth currents at Legon for the different seasons has been obtained from the following groups of months:

(1) December solstice—November, December, January, February;
(2) Equinox—March, April, September, October;
(3) June solstice—May, June, July, August.

In each case, the five quiet days of each month were used. The $S_q$-variations for the different seasons and for both the N–S and E–W components are given in Figs. 4(a), (b) and (c). A comparison of the graphs shows:

(1) The morning maximum in winter is at 0800 hours while for the equinoxes and summer it is at 1000 hours. The afternoon maximum occurs, in all cases at 1600 hours;
(2) The daily ranges in both the E–W and N–S directions are much larger for the June solstice than they are for the December solstice;
(3) In the June solstice there is an additional maximum in both E–W and N–S components at 0600 hours;
(4) After the afternoon maximum, the value of the earth currents falls to approximately zero at 2000 hours in the case of the June solstice but for the December solstice and the Equinoxes the rate of decrease is less rapid.

*Comparison of earth current values and values of the magnetic elements*

Values of the magnetic elements at Legon have only recently become available, and they show a close similarity to those at Ibadan, where magnetic recordings have already been made for a period of several years. As a result, in the first instance, the values of earth currents at Legon have been compared with the values of the horizontal and vertical component of the magnetic field at Ibadan. The closest relation has been found to be that between the earth current magnitude $R$ and the rate of change of the vertical component, $dZ/dt$, as shown in Fig. 2(a).

The diurnal variation of $dZ/dt$ at Ibadan and $R$ at Legon for quiet days and for the three seasons has also been determined and is given in Fig. 5. In these graphs, $R$ is again considered positive if the earth currents flow towards the north. It is interesting to note the following common features of the two phenomena, occurring in the December and June solstices:

(a) The daily ranges of $dZ/dt$ and of $R$ are both considerably larger during the June solstice than during the December solstice;
(b) The diurnal variation curves for $dZ/dt$ and $R$ show only two maxima in the December solstice but during the June solstice they both have an additional maximum at 0600 hours.
(c) The correspondence between the earth current and magnetic values at the Equinoxes is not so clear.

Onwumechilli and Alexander (1959) have reported that the values of $Z$ at Ibadan are 3 to 4 times greater than any others previously obtained for equatorial stations. It would seem reasonable, therefore, on account of apparent correspondence between $R$ and $dZ/dt$, that the very large values of $R$ are to some extent, at least, produced by the anomalously large values of $Z$.

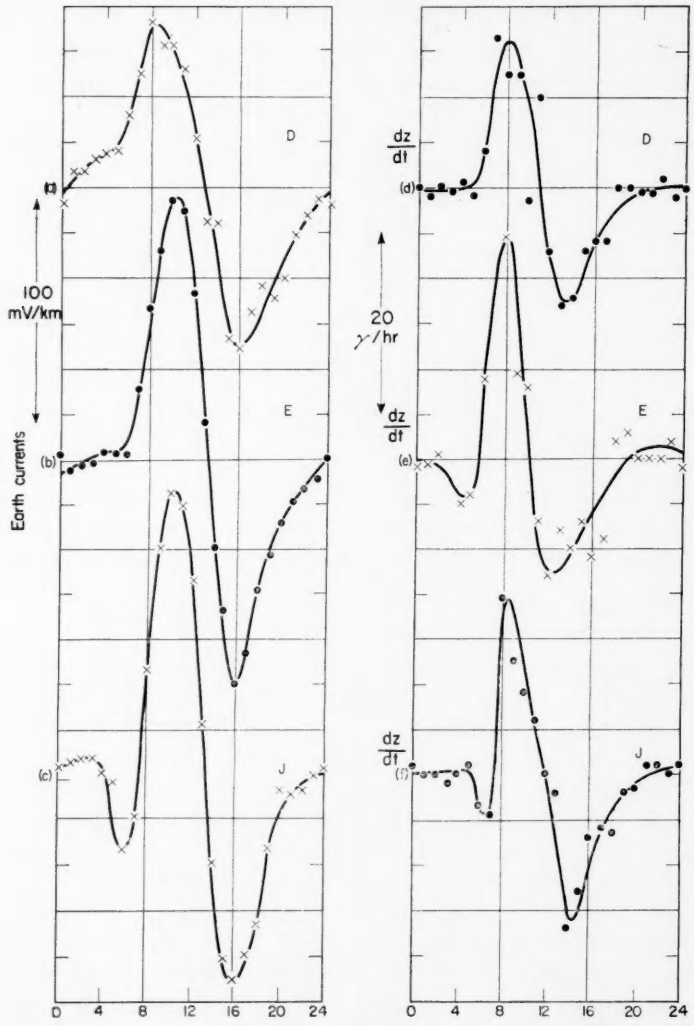


Fig. 5. $S_q$-variation for the three seasons. (a, b, c)—earth current magnitude at Legon. (d, e, f)—$dZ/dt$ at Ibadan.

*Comparison of earth current measurements at the Equatorial stations*

The only earth current data, so far reported from an equatorial station has been that from Huancayo, Peru, where the mean amplitude of the diurnal variation is of the order of 3 mV/km. In an attempt to find an explanation for the very large daily range of 180 mV/km at Legon, two factors have so far been considered:

(1) The influence on the Legon results of the nearby South Atlantic Ocean;

(2) The relation between the daily range of earth current magnitude and the daily range of $dZ/dt$.

A comparison of the earth current amplitudes at Legon with those at Tamale,

Northern Ghana and at Ibadan, Nigeria is given in Table 1, together with values of the daily range in $Z$ at these stations.

Table 1.

| Locality | Distance from coast (km) | Earth current amplitude (mV/km) | $Z$ amplitude ($\gamma$) |
|---|---|---|---|
| Legon | 11 | 180 | 70 |
| Ibadan | 112 | 150 | 55 |
| Tamale | 450 | 10 | 11 |

From the values given in this Table, it can be seen that the amplitudes of the earth currents at Legon and Ibadan are of the same order of magnitude, despite the fact that Ibadan is 101 km further from the sea. This suggests that the effect of the nearby ocean on the earth current amplitude is not large. On the other hand, a comparison of the values of earth current amplitude $R$ at the three stations with the amplitude of the $Z$ daily variation, suggests a possible dependence of earth current amplitude on the amplitude of the $Z$ element.

## Discussion

The correspondence between the earth current amplitude and the daily range of $Z$ and the similarity in shape of the diurnal variation curves for earth currents and $dZ/dt$ are both factors which suggest that earth currents near the magnetic equator are closely linked with the vertical component of the earth's magnetic field.

Large values of $dZ/dt$ occur on either side of the equatorial electrojet; Ibadan and Legon both being close to the region of maximum diurnal variation of $Z$. Tamale is almost exactly under the centre of the electrojet and hence the diurnal variation of $Z$ is small. It would be of very great interest to ascertain whether large values of earth currents are observed north of the electrojet.

*Acknowledgements*—We are grateful to the Ghana National Committee for the I.G.Y. for a grant which enabled us to carry out this work, and to Dr. K. Burkhart for his guidance in the establishment of the earth current installation. We are also grateful to Prof. N. S. Alexander and Dr. C. A. Onwumechilli for permission to use the magnetic data from the University College of Ibadan.

## References


| | | |
|---|---|---|
| Burkhart K. | 1956 | *Geofis. Pur. Appl.* **33,** 49. |
| Egedal J. | 1948 | *Nature, Lond.* **16,** 443. |
| Gish O. H. | 1936 | *J. Wash. Acad. Sci.* **26,** 267. |
| Gish O. H. and Rooney, W. J. | 1930 | *Terr. Magn. Atmos. Elect.* **35,** 213. |
| Hessler V. P. and Wescott E. M. | 1959 | *Nature, Lond.* **184,** 627. |
| Onwumechilli C. A. | 1959 | *J. Atmosph. Terr. Phys.* **13,** 222. |
| Onwumechilli C. A. and Alexander N. S. | 1959 | *J. Atmosph. Terr. Phys.* **16,** 106. |
| Yoshimatsu T. | 1957 | *Mem. Kakioka Magn. Observ.* Suppl. OK. 1. |




VOL. 20 1961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 110 to 121. Pergamon Press Ltd. Printed in Northern Ireland

# Excitation of green and red lines in the night airglow

S. N. Ghosh and A. Sharma
J. K. Institute of Applied Physics, University of Allahabad, Allahabad, India

(*Received* 15 *August* 1960)

**Abstract**—The rate of collisional deactivation of excited O atoms which leads to the emission of the $\lambda$5577 line in the night airglow has been estimated between the altitudes 50 and 200 km, after taking into account all processes by which these excited atoms may be quenched. The excitation mechanisms of this line are then considered in the light of recent rocket observations which give 100 km as the altitude of its emission. It is found that Chapman's mechanism can alone explain the observed characteristics of the $\lambda$5577 line.

The diurnal variation of the intensity of this line has been attributed to the tidal oscillations in the upper atmosphere. On the other hand, irregular temporal and spatial variations of this line are caused by turbulence.

The excitation of the $\lambda$6300 line is also considered. It is found that dissociative recombination of $O_2^+$ ions with electrons lead to the production of the red oxygen line in the night airglow.

## 1. Introduction

In order to understand properly the phenomenon of night airglow, the mechanisms for the removal of excited species which give rise to the night airglow should be clearly understood. In this paper, the excitation of oxygen lines in the night airglow has been considered after taking into account the quenching of excited atoms. Since the green $\lambda$5577 line originating from the $^1S$ level of OI atoms is the strongest feature of the night airglow spectrum, the probability of removal of OI ($^1S$) atoms at different altitudes is first considered. The excitation mechanism for the red auroral line is then taken up.

## 2. Excitation Mechanism of the $\lambda$5577 Line

Any proposed mechanism for the excitation of the $\lambda$5577 line should explain the following observed facts (Heppner and Meredith, 1958):

(i) That the $\lambda$5577 line is emitted at an altitude around 100 km, and

(ii) that the emission rate is about $2{\cdot}8 \times 10^7$ quanta $cm^{-3}\,sec^{-1}$.

Considering the composition at 100 km, where the green line is emitted (Heppner and Meredith, 1958; Elsässer and Siedentopf, 1956; Tousey, 1958), the excitation of the $\lambda$5577 line may be produced by a reaction of the following type:

(a) Bimolecular reaction involving neutral particles.

(b) Bimolecular reaction involving ions and electrons.

(c) Trimolecular reaction.

A bimolecular reaction involving neutral particles can be represented as follows:

$$X + YO = XY + O'$$

where X and Y are an atom or a molecule. From consideration of energy

$$D(XY) \geqslant D(YO) + E(O^1S)$$

where $D(XY)$ and $D(YO)$ are the dissociation energies of XY and YO molecules respectively and $E(O^1S)$ is the excitation energy of OI($^1S$) atoms.

To obtain the minimum value of $D(XY)$, it should be noted that of all diatomic oxides present at 100 km altitude, OH has the minimum dissociation energy. Substituting $D(OH) = 4{\cdot}36$ eV and $E(O^1S) = 4{\cdot}1$ eV, we have

$$D(HX) = 4{\cdot}36 + 4{\cdot}1 = 8{\cdot}46 \text{ eV}.$$

A hydride of such high dissociation energy is not known to be present at the altitude of 100 km. Thus, a bimolecular process involving neutral particles cannot be a possible mechanism for the excitation of the $\lambda$5577 line in the night airglow.

A bimolecular reaction involving ions cannot be considered as a possible process because the available energy for the reaction is much larger than the excitation energy of an $OI(^1S)$ atom, and hence the probability of the reaction is low. Also, from the low concentration of ions at 100 km, it can be shown by a simple calculation that the observed rate of emission of $\lambda$5577 cannot be obtained by this type of reaction.

Regarding the third type of reaction, it should be noted that the following reaction can only give the observed emission rate:

$$O + O + O = O_2 + O(^1S). \tag{1}$$

This is the famous Chapman's reaction. The coincidence of the altitude of maximum rate of $\lambda$5577 emission with that of the maximum oxygen atom concentration gives added support in favour of this reaction (Fig. 1).

VOL. 20 1961

### 3. Reactions Responsible for Quenching of $OI(^1S)$ Atoms

Since the $\lambda$5577 line in the night airglow originates from altitudes around 100 km, the region between 50 and 200 km has been selected for the present study. The concentration of atmospheric constituents given by Miller (1957) has been assumed. In the above region the following reactions can remove $OI(^1S)$ atoms:

$$O^* + N \rightarrow O + N^* \tag{2}$$
$$O^* + Na \rightarrow O + Na^* \tag{3}$$
$$O^* + H + M \rightarrow OH' + M \tag{4}$$
$$O^* + O + M \rightarrow O_2' + M \tag{5}$$
$$O^* + N + M \rightarrow NO' + M \tag{6}$$
$$O^* + Na + M \rightarrow NaO' + M \tag{7}$$
$$O^* + O_2(^3\Sigma_g^-) \rightarrow O + O_2(^1\Sigma_g^+, {}^1\Delta_g) \tag{8}$$
$$O^* + (H_2, N_2, NO, OH) \rightarrow O + (H_2, N_2, NO, OH, \text{vib. excited}) \tag{9}$$
$$O^* + O_2 \rightarrow O_2' + O \tag{10}$$
$$O^* + N_2 \rightarrow NO' + N \tag{11}$$
$$O^* + H_2 \rightarrow OH' + O \tag{12}$$
$$O^* + NO \rightarrow O_2' + N \tag{13}$$
$$O^* + OH \rightarrow O_2' + H \tag{14}$$
$$O^* + O_2 + M \rightarrow O_3' + M \tag{15}$$
$$O^* + N_2 + M \rightarrow N_2O' + M \tag{16}$$
$$O^* + H_2 + M \rightarrow H_2O' + M \tag{17}$$
$$O^* + OH + M \rightarrow HO_2' + M \tag{18}$$
$$O^* + NO + M \rightarrow NO_2' + M \tag{19}$$

where O* represents OI($^1S$) atoms and dash (′) signifies an excited (translationally, electronically, vibrationally or rotationally) molecule.

Since the concentration of sodium atoms in the atmosphere is small, reactions (3) and (7) involving Na atoms can be neglected. Also, since the concentration of N atoms is lower by two orders than those of $O_2$ and $N_2$, the reaction (2) can be neglected. Again, since the concentrations of reacting species at altitudes 50–200 km are very low and the cross sections for three-body reactions very low compared to those of two-body reactions, all quenching processes involving three bodies

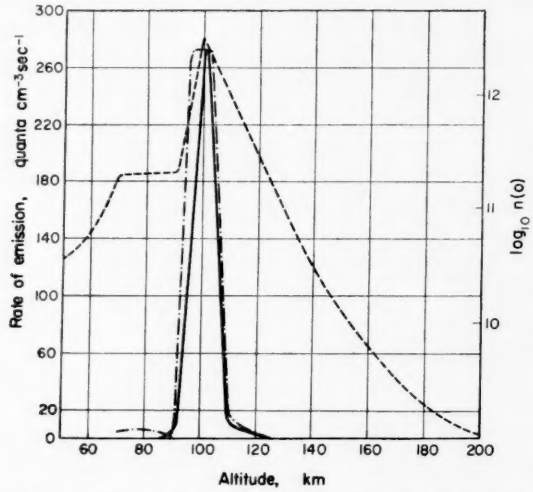


Fig. 1. Variation of rates of emission of $\lambda$5577 with altitude in the night airglow, (dash–dot curve is given by HEPPNER and MEREDITH (1958), full line curve is the calculated one). The distribution of oxygen atoms as given by MILLER (1957) is also shown in the same figure (dash curve). Note that the altitude of maximum emission of the $\lambda$5577 line coincides with that of maximum oxygen atom concentration.

namely (4), (5), (6) and (15)–(19) can be neglected. Hence, the only possible reactions for the removal of OI($^1S$) atoms are (8)–(14). Of these reactions, those involving $H_2$, OH and NO which have comparatively low concentrations can be neglected. Therefore, we finally obtain that reactions responsible for the quenching of OI($^1S$) atoms are (8)–(11). Of these reactions, (8) and (9) are of a physical quenching type, whereas reactions (10) and (11) involving exchange of atoms between reacting species are of a chemical type. (In reaction (10), it may be considered that the $O_2$ molecule is at first broken up into two O atoms, one of these atoms then combines with O* forming an $O_2'$ molecule.)

## 4. QUENCHING PROBABILITY

No precise information regarding the quenching cross-section involving metastable atoms is available. However, according to LAIDLER (1954), the coefficient of quenching reactions involving non-metastable atoms is of the order of $10^{-12}$ cm$^3$/sec. Assuming that the coefficient of quenching of metastable OI($^1S$) atoms by $N_2$ and $O_2$ molecules is one order smaller than the above value, the



VOL. 20 1961

probability of quenching of OI($^1S$) atoms $P_q$, is given by

$$P_q = [n(O_2) + n(N_2)] \times 10^{-13} \text{ sec}^{-1}$$

where $n(O_2)$ and $n(N_2)$ are the concentrations of $O_2$ and $N_2$ molecules respectively. Assuming the concentrations of atmospheric constituents given by MILLER (1957), the altitude variation of $P_q$ is computed and is given in Table 1.

Fig. 2 shows the probability of quenching of OI($^1S$) atoms between altitudes 50 and 200 km. It may be recalled that the altitude variation of the concentration

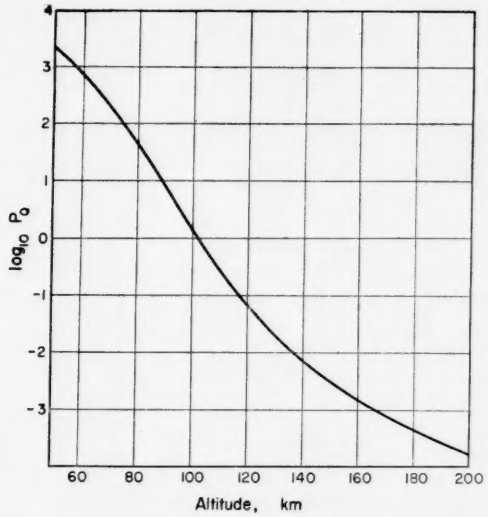


Fig. 2. Illustrating the variation of quenching probability with altitude obtained from Table 1.

of O atoms given by MILLER (1957) and that of the rate of $\lambda$5577 emission are plotted in Fig. 1. An examination of Figs. 1 and 2 reveals the following facts:

1. The altitude of the maximum rate of $\lambda$5577 emission coincides with that of the maximum oxygen atom concentration, indicating that the process for the excitation of the $\lambda$5577 line in the night airglow involves oxygen atoms in the ground state.

2. The quenching probability and the transition probability of ($^1S$–$^1D$) OI transition are of the same order of magnitude.

## 5. Rate of Excitation of the $\lambda$5577 Line

From rocket-borne experiments, HEPPNER and MEREDITH (1958) showed that the rate of emission of $\lambda$5577 at 100 km is 280 quanta cm$^{-3}$ sec$^{-1}$. Therefore

$$n(O^1S) \times A_{5577} = 280$$

where $A_{5577}$ is the probability of transition from the $^1S$ to $^1D$ state and $n(O^1S)$ is the concentration of O($^1S$) atoms at 100 km. Assuming the value of $A_{5577}$ given

by Aller (1953), namely 1·28 $sec^{-1}$ we have

$$n(O^1S) = \frac{280}{1{\cdot}28} = 218{\cdot}8 \text{ atom cm}^{-3}.$$

From Table 1, the probability of quenching of OI($^1S$) atoms at 100 km is 1·3 $sec^{-1}$. Therefore, the rate of quenching of OI($^1S$) atoms is given by

$$218{\cdot}8 \times 1{\cdot}3 = 284{\cdot}4 \text{ cm}^{-3} \text{ sec}^{-1}.$$

Hence, the total rate of de-excitation is given by

$$280 + 284{\cdot}4 = 564{\cdot}4 \text{ cm}^{-3} \text{ sec}^{-1}$$

which at equilibrium condition is the rate of production of OI($^1S$) atoms.

If the rate coefficient of Chapman's process is $k_1$, then at equilibrium condition

$$n(O)^3_{100} \times k_1 = 564{\cdot}4$$

where $n(O)_{100}$ is the concentration of oxygen atoms at 100 km. Substituting the value of $n(O)$ from Table 1, we obtain

$$k_1 = \frac{564{\cdot}4}{3{\cdot}4} \times 10^{-36} \text{ cm}^6 \text{ sec}^{-1}$$

$$= 1{\cdot}66 \times 10^{-34} \text{ cm}^6 \text{ sec}^{-1}.$$

This is a plausible value of $k_1$ for a three-body reaction.

To determine the variation of emission rate with altitude, let us consider the equilibrium condition, when

$$n(O)^3 \,.\, k_1 = n(O^1S) \,.\, [A_{5577} + n(M) \,.\, k_3]$$

where $k_3$ is the rate coefficient of the quenching reaction discussed in Section 3. Since the effective quenchers are $N_2$ and $O_2$ molecules, $n(M)$ is equal to the sum of the concentrations of $N_2$ and $O_2$ molecules. Therefore,

$$n(O^1S) = \frac{n(O)^3 \,.\, k_1}{A_{5577} + n(M) \,.\, k_3}$$

and the emission rate

$$r_{5577} = n(O^1S) \,.\, A_{5577}$$

$$= \frac{n(O)^3 \,.\, k_1}{A_{5577} + n(M) \,.\, k_3} A_{5577}.$$

Substituting the values of $k_1$ and $k_3$ and the values of $n(O)$ and $n(M)$ at different altitudes from Table 1, the variation of the emission rate with altitude is computed and is plotted in Fig. 2. It can be seen from this figure that the agreement between the theoretical and experimental values is good.



VOL. 20 1961

Table 1. Calculated variation of the probability of quenching of OI($^1S$) atoms, $P_q$, with altitude

| Altitude (km) | $n$(O) ($cm^{-3}$) | $n(O_2)$ ($cm^{-3}$) | $n(N_2)$ ($cm^{-3}$) | $P_q$ ($sec^{-1}$) |
|---|---|---|---|---|
| 50 | $3{\cdot}45 \times 10^{10}$ | $4{\cdot}72 \times 10^{15}$ | $1{\cdot}76 \times 10^{16}$ | $2{\cdot}23 \times 10^{3}$ |
| 60 | $6{\cdot}42 \times 10^{10}$ | $1{\cdot}52 \times 10^{15}$ | $3{\cdot}67 \times 10^{15}$ | $5{\cdot}19 \times 10^{2}$ |
| 70 | $1{\cdot}90 \times 10^{11}$ | $4{\cdot}37 \times 10^{14}$ | $1{\cdot}63 \times 10^{15}$ | $2{\cdot}06 \times 10^{2}$ |
| 80 | $2{\cdot}15 \times 10^{11}$ | $9{\cdot}43 \times 10^{13}$ | $3{\cdot}52 \times 10^{14}$ | $4{\cdot}46 \times 10$ |
| 90 | $2{\cdot}30 \times 10^{11}$ | $1{\cdot}74 \times 10^{13}$ | $6{\cdot}49 \times 10^{13}$ | 8·23 |
| 100 | $3{\cdot}40 \times 10^{12}$ | $1{\cdot}40 \times 10^{12}$ | $1{\cdot}16 \times 10^{13}$ | 1·30 |
| 110 | $1{\cdot}06 \times 10^{12}$ | $1{\cdot}62 \times 10^{11}$ | $2{\cdot}58 \times 10^{12}$ | $2{\cdot}74 \times 10^{-1}$ |
| 120 | $3{\cdot}13 \times 10^{11}$ | $3{\cdot}00 \times 10^{10}$ | $6{\cdot}95 \times 10^{11}$ | $7{\cdot}25 \times 10^{-2}$ |
| 130 | $9{\cdot}46 \times 10^{10}$ | $9{\cdot}08 \times 10^{9}$ | $2{\cdot}10 \times 10^{11}$ | $2{\cdot}19 \times 10^{-2}$ |
| 140 | $3{\cdot}12 \times 10^{10}$ | $2{\cdot}62 \times 10^{9}$ | $6{\cdot}94 \times 10^{10}$ | $7{\cdot}20 \times 10^{-3}$ |
| 150 | $1{\cdot}29 \times 10^{10}$ | $8{\cdot}61 \times 10^{8}$ | $2{\cdot}87 \times 10^{10}$ | $2{\cdot}95 \times 10^{-3}$ |
| 160 | $6{\cdot}19 \times 10^{9}$ | $3{\cdot}41 \times 10^{8}$ | $1{\cdot}38 \times 10^{10}$ | $1{\cdot}41 \times 10^{-3}$ |
| 170 | $3{\cdot}30 \times 10^{9}$ | $1{\cdot}54 \times 10^{8}$ | $7{\cdot}33 \times 10^{9}$ | $7{\cdot}48 \times 10^{-4}$ |
| 180 | $1{\cdot}90 \times 10^{9}$ | $7{\cdot}73 \times 10^{7}$ | $4{\cdot}23 \times 10^{9}$ | $4{\cdot}30 \times 10^{-4}$ |
| 190 | $1{\cdot}37 \times 10^{9}$ | $4{\cdot}32 \times 10^{7}$ | $2{\cdot}52 \times 10^{9}$ | $2{\cdot}56 \times 10^{-4}$ |
| 200 | $1{\cdot}01 \times 10^{9}$ | $2{\cdot}51 \times 10^{7}$ | $1{\cdot}55 \times 10^{9}$ | $1{\cdot}57 \times 10^{-4}$ |

## 6. Diurnal Variation of the Intensity of the λ5577 Line

The diurnal variation of this line in the night airglow has been thoroughly investigated (Barbier *et al.*, 1954; Roach, 1954a, b) and has been found to possess the following characteristics:

i. The diurnal variation of the emission rate at the zenith is irregular. Maximum emission rate occurs between 2·5 hr before and after local midnight. Sometimes more than one maximum is observed.

ii. Simultaneous observation of the intensity of the λ5577 line at two places shows that the *general* feature of the excitation pattern† depends on the local time. On the other hand, the *detailed* features of the excitation pattern for the same local time vary irregularly from place to place. Often the excitation pattern is found to have a motion.

The lack of information of the diurnal variation of the physical state (temperature, pressure, density, etc.) of the upper atmosphere at the emitting layer prevents us giving a quantitative explanation of the diurnal variation of the λ5577 line emission rate. Therefore, instead of attempting to explain the observed variation of the λ5577 line, the required variation of the density of atmospheric gases to produce the observed variation of this line intensity has been calculated.

In Section 5, it has been shown that the rate of emission of the λ5577 line is given by

$$r_{5577} = \frac{n(\mathrm{O})^3 \,.\, k_1 \,.\, A_{5577}}{A_{5577} + n(\mathrm{M}) \,.\, k_3} \,.$$

† It has been shown that the intensity of the λ5577 line in the night airglow is not uniform over the whole of the emitting layer, but has a distribution called the excitation pattern. (The curves passing through the points of equal intensity are called isophotes, so that the excitation pattern of the emitting layer can be represented by an isophote map.)

Therefore, the number of photons emitted per cm² column per sec is given by

$$R_{5577} = \frac{H \,.\, n(\mathrm{O})^3 \,.\, k_1 A_{5577}}{A_{5577} + n(\mathrm{M}) \,.\, k_3}$$

where $H$ is the thickness of the emitting layer.

It has been assumed that the density and thickness of the emitting layer undergo periodic change due to tidal oscillations† of the upper atmosphere while the total number of particles in the emitting layer remains constant. Also, remembering that the emitting layer is thin, we can write

$$n(\mathrm{O}) \times H = \text{Constant} = N(\mathrm{O}) \text{ atoms/cm}^2 \text{ column} \qquad \text{(A)}$$

$$n(\mathrm{M}) \times H = \text{Constant} = N(\mathrm{M}) \text{ particles/cm}^2 \text{ column} \qquad \text{(B)}$$

The emission rate can now be written as

$$R_{5577} = \frac{N(\mathrm{O})^3 \,.\, k_1 \,.\, A_{5577}}{H[A_{5577} \,.\, H + N(\mathrm{M}) \,.\, k_3]} \text{ photon/cm}^2 \text{ column.}$$

The rate of change of $R_{5577}$ with the thickness of the layer is given by

$$\frac{\Delta R_{5577}}{\Delta H} = -R_{5577} \frac{2H\, A_{5577} + N(\mathrm{M}) \,.\, k_3}{H[A_{5577} \,.\, H + N(\mathrm{M}) \,.\, k_3]}$$

or

$$\frac{\Delta R_{5577}}{R_{5577}} = -\frac{2H \,.\, A_{5577} + N(\mathrm{M}) \,.\, k_3}{A_{5577} + N(\mathrm{M}) \,.\, k_3} \times \frac{\Delta H}{H} \,.$$

Assuming the emitting layer to be 10 km thick and inserting the numerical values from Table 1, we have

$$\frac{\Delta R_{5577}}{R_{5577}} = -1{\cdot}89 \times \frac{\Delta H}{H} \,.$$

Using the relation (B), we obtain

$$\frac{\Delta H}{H} = -\frac{\Delta n(\mathrm{M})}{n(\mathrm{M})}$$

or

$$\frac{\Delta R_{5577}}{R_{5577}} = 1{\cdot}89 \times \frac{\Delta n(\mathrm{M})}{n(\mathrm{M})} \,.$$

Therefore

$$\frac{\Delta n(\mathrm{M})}{n(\mathrm{M})} = 0{\cdot}53 \times \frac{\Delta R_{5577}}{R_{5577}} \,.$$

From the mean diurnal variation of the emission rate of the λ5577 line obtained by Roach (1954b), it can be shown that $\Delta R_{5577}/R_{5577}$ is 0·43. Hence

$$\frac{\Delta n(\mathrm{M})}{n(\mathrm{M})} = 0{\cdot}23.$$

Therefore, in order to explain the observed diurnal variation of the green OI line

† The nocturnal variation in the density of atmospheric gases may be due to tidal oscillations in the upper atmosphere produced by solar and lunar pull and the periodic heating by the sun.



VOL. 20 1961

intensity in the night airglow, the diurnal variation (for a period of 10 hr) of the total particle density should be 23 per cent. Also, it can be concluded that or since the emission rate of the $\lambda$5577 line is at a maximum between 2·5 hr before after local midnight, the maximum total particle density at the emitting layer should also occur during the same period.

Contrary to the above conclusion, LISZKA (1956) holds that the maximum emission rate of the $\lambda$5577 line occurs when the density of atmospheric gases at the emitting layers is at a minimum, because with the increase of the density of atmospheric gases the quenching increases thereby decreasing the emission rate. However, LISZKA has not considered the effect of density variation on the excitation mechanism of the $\lambda$5577 line. Again, although maximum rarefaction of the density at the ground occurs at the lower culmination of the moon, it does not necessarily follow that the density at the emitting layer becomes minimum at the same time.

Also, it may be remarked that the variation of atmospheric density produced by tidal oscillations is a function of local time. Therefore, from the above it follows that the emission rate of the green OI line should also depend on the local time, which is confirmed experimentally.

A tentative explanation of the detailed features of the excitation pattern can be given by assuming the presence of turbulence at the altitude of emission. (Its presence has been shown by BOOKER (1956) from observations of the fading of radar echoes from meteor trails.) Therefore, the total particle concentration per column is not uniform throughout the emitting layer and causes non-uniformity in emission rate. Moreover, the turbulence produces irregular variation of the emission rate at different places and different times.

## 7. EXCITATION OF RED OXYGEN LINES

After emitting the green $\lambda$5577 line, an oxygen atom arrives at the $^1D$ state. A transition from the former state to the ground $^3P$ level results in the emission of the red $\lambda$6300 line. Therefore, after emitting the $\lambda$5577 line by radiative transition, an oxygen atom can emit the red auroral line. However, to explain the relative intensity of green and red lines observed in the night airglow, ROACH (1954b) showed that one has to consider the production of OI($^1D$) atoms by some other mechanism in addition to those produced by radiative transition from the $^1S$ state. A search for such a mechanism can be made in the light of the following observations:

1. *Height of emission of the red OI line.* From rocket measurements of the altitude of night airglow emission, HEPPNER and MEREDITH (1958) concluded that the emission of OI red lines in the night airglow originates primarily above 160 km although they do not exclude the possibility of emission at lower altitudes, which if it occurs could be only a fraction of the total emission.

2. *Rate of emission.* According to ROACH's (1954b) observations, the average rate of emission of the red oxygen line in the night airglow is $15 \times 10^6$ quanta/cm$^2$ column.

Assuming Chapman's mechanism for the excitation of the OI red line in the night airglow, namely,

$$O + O + O = O_2 + O(^1D) \qquad (20)$$

we can calculate the rate of emission of red lines at different altitudes between 50 and 200 km in the following manner. At equilibrium condition

$$n(O^1S) \cdot A_{5577} + n(O)^3 \cdot k_1' = n(O^1D) \cdot (A_{6300} + P_{q6300}) \tag{21}$$

where

$n(O^1S)$ = concentration of OI($^1S$) atoms.

$n(O^1D)$ = concentration of OI($^1D$) atoms.

$A_{6300}$ = transition probability of the red OI line.

$P_{q6300}$ = quenching probability of OI($^1D$) atoms.

$k_1'$ = rate coefficient of reaction (20).

The rate of emission of red oxygen line is given by

$$r_{6300} = n(O^1D) \cdot A_{6300}.$$

Substituting the value of $n(O^1D)$ from (21), we have

$$r_{6300} = \frac{A_{6300}}{A_{6300} + P_{q6300}} \times \left[ n(O)^3 \cdot k_1' + \frac{n(O)^3 \cdot k_1 \cdot A_{5577}}{A_{5577} + P_{q5577}} \right]$$

and assuming $k_1' = \alpha k_1$, $P_{q6300} = \beta P_{q5577}$, we have

$$r_{6300} = \frac{n(O)^3 \cdot k_1 \cdot A_{6300}}{A_{6300} + \beta P_{q5577}} \times \left[ \alpha + \frac{A_{5577}}{A_{5577} + P_{q5577}} \right]$$

where $\alpha$ and $\beta$ are constants. Using this formula the relative variation of $r_{6300}$ with altitude can be obtained. However, since the values of $\alpha$ and $\beta$ are not precisely known and the value of $k_1$ is uncertain, a quantitative estimate of the variation of $r_{6300}$ with altitude cannot be made. A rough estimate of this variation can be made in the following way. Depending on the values of $P_{q6300}$ and $P_{q5577}$ as estimated by SEATON (1958), the upper region of the atmosphere is divided into three regions.

I. *Lower region*. This region is situated below 100 km altitude, where the following condition exists:

$$P_{q5577} \gg A_{5577}, \text{ and } P_{q6300} \gg A_{6300}.$$

In this region, the rate of emission of the red OI triplet is given by

$$r_{6300} = \frac{n(O)^3 \cdot k_1 \cdot A_{6300}}{\beta P_{q5577}} \times \left( \alpha + \frac{A_{5577}}{P_{q5577}} \right)$$

where the oxygen atom concentration increases and $P_{q5577}$ decreases with the altitude (see Fig. 1). Therefore, $r_{6300}$ rapidly increases up to 100 km altitude.

II. *Middle region*. This region is situated above 100 km and extends to 200 km, where the following condition holds, namely

$$P_{q5577} \ll A_{5577} \text{ and } P_{q6300} \gg A_{6300}.$$



VOL. 20 1961

Consequently, $r_{6300}$ is given by

$$r_{6300} = \frac{n(\mathrm{O})^3 \,.\, k_1 \,.\, A_{6300}}{\beta P_{q5577}} \times (\alpha + 1).$$

In this region the altitude variation of $r_{6300}$ is given by

$$\frac{\Delta r_{6300}}{\Delta h} = r_{6300} \times \left[\frac{3 \,.\, \Delta n(\mathrm{O})}{n(\mathrm{O}) \,.\, \Delta h} - \frac{\Delta n(\mathrm{M})}{n(\mathrm{M}) \,.\, \Delta h}\right]$$

where $h$ represents the altitude of emission and $n(\mathrm{M})$ the sum of concentration of $O_2$ and $N_2$ molecules. From the distributions of O, $O_2$ and $N_2$ given by MILLER (1957), it is found from the above relation that in this region $\Delta r_{6300}/\Delta h$ is negative, i.e. the rate of emission decreases with altitude.

III. *Upper region.* This region is located above 200 km. In this region the following conditions exist:

$$P_{q5577} \ll A_{5577}, \text{ and } P_{q6300} \ll A_{6300}$$

and therefore the rate of emission of the $\lambda$6300 line is given by

$$r_{6300} = n(\mathrm{O})^3 \,.\, k_1 \,.\, (\alpha + 1).$$

Since in this region the concentration of oxygen atoms decreases with altitude, the rate of emission also decreases. Therefore, it can be concluded that if the emission of the $\lambda$6300 line in the night airglow is also due to the Chapman process, the maximum rate of its emission should occur at about 100 km where the atomic oxygen concentration is a maximum. As this is contrary to observations, we may conclude that the process responsible for the emission of the $\lambda$6300 line is different from that by which the green line is emitted. This conclusion is supported by the fact that the seasonal and diurnal variations of the intensity of the green line are markedly different from those of the red OI line (ROACH, 1954b).

Considering the other processes mentioned in Section 2, it should be noted that at altitudes higher than 160 km, the rate of any tri-molecular process is too low to explain the observed emission.

If a bimolecular process involving neutral particles could emit the $\lambda$6300 line, it should be of the following type:

$$\mathrm{X} + \mathrm{YO} = \mathrm{XY} + \mathrm{O}(^1D).$$

For this process to be valid, the following condition should hold (see Section 2):

$$D(\mathrm{XY}) - D(\mathrm{YO}) > E(\mathrm{O}^1D)$$

where $E(\mathrm{O}^1D)$ is the excitation energy of $\mathrm{O}^1D$ atoms, which is 1·96 eV. Above 160 km, the only possible reaction† of the above type is given by

$$\mathrm{N} + \mathrm{NO} = \mathrm{N_2} + \mathrm{O}(^1D) \qquad (22)$$

† It should be noted that the above reaction cannot produce $\mathrm{O}^1S$ atoms, because the energy released is lesser than the excitation energy of $\mathrm{O}^1S$ atoms (4·2 eV).

because

$$D(N_2)\,(9{\cdot}76\text{ eV}) - D(NO)\,(6{\cdot}48\text{ eV}) = 3{\cdot}28\text{ eV} > E(O^1D).$$

At equilibrium condition, we have

$$n(N)\,.\,n(NO)\,.\,k_4 = n(O^1D)\,.\,(A_{6300} + P_{q6300})$$

where $n(N)$, $n(NO)$, and $n(O^1D)$ are respectively the concentrations of N, NO and $O^1D$ and $k_4$ is the rate coefficient of reaction (22). Therefore, according to the above process, the rate of emission of the red oxygen line in the night airglow is given by

$$r_{6300} = n(O^1D)\,.\,A_{6300} = \frac{n(N)\,.\,n(NO)\,.\,k_4}{A_{6300} + P_{q6300}} \times A_{6300}.$$

In the region around 100 km, the magnitude of $P_{q6300}$ is much larger than $A_{6300}$, and therefore $r_{6300}$ is given by

$$r_{6300} = \frac{n(N)\,.\,n(NO)\,.\,k_4}{P_{q6300}} \times A_{6300}.$$

Using the distribution of N and NO given by MILLER (1957) and the value of $P_{q6300}$ given by SEATON (1958), one finds that the $\lambda$6300 line should be emitted at about 100 km (contrary to observation) and that the rate of emission be decreased with altitude, which is also contrary to observations. Hence, it can be concluded that the above process is not responsible for the emission of the red oxygen line in the night airglow. It may contribute a small fraction of this emission at lower altitudes as reported by HEPPNER and MEREDITH (1958).

It is to be noted that because of the low concentration of negative ions, the $\lambda$6300 line cannot be emitted by mutual neutralization of positive and negative ions. Also, radiative recombination between $O^+$ and $e$ is not a plausible reaction as its coefficient (BATES *et al.*, 1939) and the concentrations of ions and electrons are very low.

As regards the dissociative recombination process, it may be noted that BATES (1956) has shown that the following reaction can only account for the observed high recombination coefficient in the $F$-layer:

$$O_2^+ + e = O' + O'$$

where dash signifies excited atoms. Assuming that either one or both oxygen atoms are excited to $^1D$ level, one finds that at equilibrium condition, the following relation holds:

$$b\,.\,n(O_2^+)\,.\,n(e)\,.\,\alpha = n(O^1D)(A_{6300} + P_{q6300})$$

where $\alpha$ is the rate coefficient of the above process, and $b$ may have a value of 1 or 2 depending on whether one or both atoms are excited. The rate of excitation of the red OI line is then given by

$$r_{6300} = \frac{b\,.\,n(O_2^+)\,.\,n(e)\alpha}{A_{6300} + P_{q6300}} \times A_{6300}.$$



VOL. 20 1961

Since, in the region above 160 km

$$P_{q6300} \ll A_{6300}$$

we have

$$r_{6300} = bn(O_2^+) \, . \, n(e) \, . \, \alpha.$$

Substituting the value of $\alpha$ which is of the order of $10^{-7}$ cm$^3$ sec$^{-1}$ (BATES, 1950) and the value of the electron density given by JACKSON and SEDDON (1958) and putting $b = 1$ in the above equation, it is found that the observed rate of emission can be explained if the concentration of $O_2^+$ is $5 \times 10^3$ ions cm$^{-3}$. This is the expected (BATES, 1956) concentration of $O_2^+$ ions in the $F$-region of the ionosphere. It may be noted that by assuming the above process, CHAMBERLAIN (1958) explained the observed rate of decrease of $r_{6300}$ after sunset at the altitude of emitting layer. Moreover, BARBIER (1957) has shown strong correlation between the intensity of the $\lambda$6300 line in the night airglow with the critical frequency and the height for vertical reflexion of radio waves from the $F$-region of the ionosphere. Also, ST. AMAND (1955) has established a definite correlation between the intensity of the $\lambda$6300 line with the $F$-layer ionization. Hence, it can be concluded that the most probable mechanism for the excitation of the red oxygen line in the night airglow is the dissociative recombination of molecular oxygen ions with electrons.

## REFERENCES


| | | |
|---|---|---|
| ALLER L. H. | 1953 | *Astrophysics.* Ronald Press, New York. |
| BARBIER D. | 1957 | *C.R. Acad. Sci., Paris* **244,** 2077. |
| BARBIER D., DUFAY J. and WILLIAMS D. R. | 1954 | *Proceedings of the Conference on Auroral Physics*, p. 137. Geophysical Research Paper No. 30, AF Cambridge Research Centre, Massachusetts. |
| BATES D. R. | 1950 | *Phys. Rev.* **77,** 718; **78,** 492. |
| BATES D. R. | 1956 | *Solar Eclipse and the Ionosphere*, p. 197. Pergamon Press, London. |
| BATES D. R., BUCKINGHAM R. A., MASSEY H. S. W. and UNWIN J. J. | 1939 | *Proc. Roy. Soc.* A **170,** 320. |
| BOOKER H. G. | 1956 | *J. Geophys. Res.* **61,** 673. |
| CHAMBERLAIN J. W. | 1958 | *Astrophys. J.* **127,** 54. |
| ELSÄSSER H. and SIEDENTOPF H. | 1956 | *J. Atmosph. Terr. Phys.* **8,** 222. |
| HEPPNER J. P. and MEREDITH L. H. | 1958 | *J. Geophys. Res.* **63,** 51. |
| JACKSON J. E. and SEDDON J. C. | 1958 | *J. Geophys. Res.* **63,** 197. |
| LAIDLER K. J. | 1954 | *Chemical Kinetics of Excited States*, p. 159. Oxford University Press. |
| LISZKA L. | 1956 | *Acta Phys. Polon.* **15,** 305. |
| MILLER L. E. | 1957 | *J. Geophys. Res.* **62,** 351. |
| ROACH F. E. | 1954a | *Proceedings of the Conference on Auroral Physics*, p. 185. Geophysical Research Paper No. 30, AF Cambridge Research Centre, Massachusetts. |
| ROACH F. E. | 1954b | *Proc. Nat. Acad. Sci., Wash.* **40,** 950. |
| ROACH F. E. | 1955 | *Ann. Géophys.* **11,** 1. |
| ST. AMAND P. | 1955 | *Ann. Géophys.* **11,** 450. |
| SEATON M. J. | 1958 | *Astrophys. J.* **127,** 67. |
| TOUSEY R. | 1958 | *Ann. Géophys.* **14,** 186. |

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 122 to 130. Pergamon Press, Ltd. Printed in Northern Ireland

# A theoretical model of electron density distribution along a geomagnetic line of force in the exosphere

R. L. Dowden
Ionospheric Prediction Service, Hobart, Tasmania, and the University of Tasmania

(*Received* 17 *August* 1960)

**Abstract**—By consideration of charged particles spiralling in geomagnetic tubes of force, particle densities are derived for a source of particles filling the tubes at (i) the equatorial plane or (ii) the earth's surface (ionosphere). The first case gives a distribution of roughly constant density for geomagnetic latitudes ($\lambda$) up to 60° and the second to one of density roughly proportional to magnetic field strength for $\lambda < 60°$ and $R < 1{\cdot}25$ earth's radii.

## 1. Introduction

Over the last few years whistler and micropulsation studies have yielded information about electron densities in the exosphere out to several earth's radii. So far this information has not been sufficient to determine whether or not the electron density varies with latitude. In studies of propagation of radio and hydromagnetic waves along the geomagnetic lines of force and of the generation of v.l.f. radio noise, dawn chorus, etc., it is necessary to assume a distribution of electron density along the line of force. The assumption usually made is that electron density is distributed with spherical symmetry about the centre of the earth. This would be logical in the absence of a magnetic field as any local discontinuities would tend to diffuse over a sphere. In the presence of a magnetic field, however, charged particles are mainly confined to a tube of force so that diffusion would take place along the magnetic field lines.

This paper is an attempt to take this effect into account and derive the distribution along a line of force on the assumption that the tubes of force are supplied with charged particles (electrons, protons, etc.) from near the earth's surface (e.g. the upper ionosphere). The particles will move in helical orbits along the line of force. Since the cross-section of the tube of force rapidly increases towards the equatorial plane and since the pitch of the helices decreases (i.e. opens out) in this direction we would expect the density to decrease. Collisions are neglected except in the source regions near the earth's surface. Here down-coming particles which are "lost" are replaced (assuming equilibrium) by an equal number of up-going ones.

The density distribution of the faster particles which produce v.l.f. radio noise and possibly aurora during magnetic storms is also derived on the assumption that these particles are injected into the tubes in regions farthest from the earth, that is near the equatorial plane.

Finally the electron density at any point in space is derived on the rough assumption that the electron density in the source regions (ionosphere) varies with geomagnetic latitude ($\lambda$) as cos $\lambda$.

In all cases symmetry about the geomagnetic axis is assumed. Hence the relatively slow drift of charged particles in longitude has been neglected.



VOL.
20
1961

## 2. Theory

The various equations describing the dipole magnetic field, a line of force and a tube of force are given (Alfvén, 1950; Chapman and Sugiura, 1956):

$$R = R_0 \cos^2 l$$

$$R_0 = \sec^2 \lambda$$

$$H = \eta H_0$$

$$HA = H_0 A_0 \text{ hence } A = A_0/\eta$$

$$\frac{ds}{dl} = R_0 \phi \cos l$$

where:

$R$ = radius vector in earth radii
$l$ = latitude angle
} polar coordinates

$\lambda$ = geomagnetic latitude at which the line of force cuts the earth's surface
$H$ = intensity of the earth's magnetic field
$\eta = \phi \sec^6 l$
$\phi = (1 + 3 \sin^2 l)^{1/2}$
$A$ = cross section area of tube of force
$s$ = distance along line of force

Subscript "0" denotes values for which $l = 0$ (i.e. in equatorial plane) and subscript "$\lambda$" values for which $l = \lambda$ (earth's surface). We note immediately that:

$$\eta_0 = \phi_0 = 1.$$

We consider one of the particles spiralling along the line of force determined by $\lambda$. At the point determined by "$l$", let its pitch angle be $\psi$ and its speed $v$. Equations governing its motion (Gold, 1959) are:

$$v = v_0 \text{ (constant)}$$

$$H/\sin^2 \psi = H_0/\sin^2 \psi_0$$

Hence $\sin^2 \psi = \eta \sin^2 \psi_0$.

The velocity of the "guiding centre" of the spiralling particle is:

$$\frac{ds}{dt} = v \cos \psi = v(1 - \sin^2 \psi)^{1/2}$$

$$= v(1 - \eta \sin^2 \psi_0)^{1/2}.$$

The probability of finding this particle at position "$l$" is proportional to $dt/dl$. Now,

$$\frac{dt}{dl} = \frac{dt}{ds} \cdot \frac{ds}{dl} = R_0 \phi \cos l [v(1 - \eta \sin^2 \psi_0)^{1/2}]^{-1}.$$

If we consider a large number of such particles ($l$, $\psi_0$, $v$) the density (cm$^{-3}$) will be given by

$$n(\psi_0, v) \propto \frac{dt}{dl}/A = \frac{KR_0}{A_0 v} \cdot \eta\phi \cos l \,(1 - \eta \sin^2 \psi_0)^{-1/2}$$

where $K$ is a constant of proportionality.

In the equatorial plane:

$$n_0(\psi_0, v) = \frac{KR_0}{A_0 v}(1 - \sin^2 \psi_0)^{-1/2} = \frac{KR_0}{A_0 v \cos \psi_0}.$$

Hence

$$n(\psi_0) = n_0(\psi_0)\eta\phi \cos l \cos \psi_0 (1 - \eta \sin^2 \psi_0)^{-1/2}. \tag{1}$$

We first consider the simpler case of a source of particles in the equatorial plane. If the source is isotropic (no preferred particle direction) simple geometrical considerations show that the distribution of pitch angles will be given by

$$dN_0(\psi_0) = N_0 \sin \psi_0 \, d\psi_0 \tag{2}$$

where the expression on the left hand side represents the density (at $l = 0$) of particles having pitch angles in the range $\psi_0$ and $\psi_0 + d\psi_0$. Substituting this for $n_0(\psi_0)$ in (1):

$$dN(\psi_0) = N_0\eta\phi \cos l \sin \psi_0 \cos \psi_0 (1 - \eta \sin^2 \psi_0)^{-1/2} \, d\psi_0.$$

The total density at point $l$, $\lambda$ is given by integration over all allowable values of $\psi_0$ from $\psi_0 = 0$, to $\psi_0 =$ its maximum value $\Psi_0$ given by $\eta \sin^2 \Psi_0 = 1$. Particles which had pitch $\Psi_0$ will have pitch $\psi = \pi/2$ at this point (magnetic mirror point). Particles of greater pitch will have been removed by magnetic reflection before reaching this point.

Hence

$$N = N_0\eta\phi \cos l \int_0^{\Psi_0} \sin \psi_0 \cos \psi_0 (1 - \eta \sin^2 \psi_0)^{-1/2} \, d\psi_0.$$

Putting $x^2 = 1 - \eta \sin^2 \psi_0$,

$$N = N_0\eta\phi \cos l \frac{l}{\eta} \int_0^1 dx = N_0\phi \cos l. \tag{3}$$

The expression $\phi \cos l$ is unity at $l = 0$ and at $\sin^2 l = 2/3$ ($l \sim 55°$). It has its maximum value (1·15) for $\sin^2 l = 1/3$ ($l \sim 35\frac{1}{2}°$). Simple computation shows it to be within 15 per cent of unity for $l < 60°$.

Hence this distribution which probably applies to the case of solar particles injected into the tube of force during magnetic storms gives a practically constant density along a line of force for all lines ($\lambda$) up to $\lambda = 60°$.

The thermal particles which make up the background medium are probably supplied where ionization is taking place, i.e. in the ionosphere. Accordingly we



VOL. 20 1961

consider an isotropic source at $l = \lambda$. The distribution of pitch angles is given [as in (2)] by

$$dN_\lambda(\psi_\lambda) = N_\lambda \sin \psi_\lambda \, d\psi_\lambda. \tag{4}$$

We relabel these same particles by the pitch angles ($\psi_0$) they would have on reaching the equatorial plane, where

$$\sin^2 \psi_\lambda = \eta_\lambda \sin^2 \psi_0. \tag{5}$$

Differentiating:

$$2 \sin \psi_\lambda \cos \psi_\lambda \, d\psi_\lambda = 2\eta_\lambda \sin \psi_0 \cos \psi_0 \, d\psi_0.$$

Hence

$$\sin \psi_\lambda \, d\psi_\lambda = \eta_\lambda \frac{\sin \psi_0 \cos \psi_0}{\cos \psi_\lambda} \, d\psi_0$$

$$= \eta_\lambda \sin \psi_0 \cos \psi_0 (1 - \eta_\lambda \sin^2 \psi_0)^{-1/2} \, d\psi_0.$$

Relabelling (4) and substituting:

$$dN_\lambda(\psi_0) = N_\lambda \eta_\lambda \sin \psi_0 \cos \psi_0 (1 - \eta_\lambda \sin^2 \psi_0)^{-1/2} \, d\psi_0. \tag{6}$$

From (1) we have:

$$n_0(\psi_0) = n_\lambda(\psi_0) \frac{(1 - \eta_\lambda \sin^2 \psi_0)^{1/2}}{\phi_\lambda \eta_\lambda \cos \psi_0 \cos \lambda}.$$

Replacing $n_0(\psi_0)$ by $dN_0(\psi_0)$ and $n_\lambda(\psi_0)$ by $dN_\lambda(\psi_0)$ as given in (6):

$$dN_0(\psi_0) = \frac{N_\lambda \sin \psi_0}{\phi_\lambda \cos \lambda} \, d\psi_0.$$

Integrating from $\psi_0 = 0$ to $\psi_0 = \Psi_0$ where $\sin^2 \Psi_0 = 1/\eta_\lambda$ [hence $\cos \Psi_0 = (1 - 1/\eta_\lambda)^{1/2}$]:

$$N_0 = \frac{N_\lambda}{\phi_\lambda \cos \lambda} \int_0^{\Psi_0} \sin \psi_0 \, d\psi_0$$

$$= \frac{N_\lambda}{\phi_\lambda \cos \lambda} \left[ 1 - \left( 1 - \frac{1}{\eta_\lambda} \right)^{1/2} \right]$$

$$= \frac{N_\lambda}{\phi_\lambda \cos \lambda} \zeta \left( \frac{1}{\eta_\lambda} \right). \tag{7}$$

Hence

$$dN_0(\psi_0) = \frac{N_0}{\zeta\left(\frac{1}{\eta_\lambda}\right)} \sin \psi_0 \, d\psi_0. \tag{8}$$

We can now find $N$ at any point "$l$" as in the previous argument by replacing the pitch distribution given in (2) by that in (8) except that the upper limit is now given by $\sin^2 \Psi_0 = 1/\eta_\lambda$.

That is

$$N = \frac{N_0}{\zeta(1/\eta_\lambda)} \phi \cos l \, [(1 - \eta \sin^2 \psi_0)^{1/2}]_{\Psi_0}^0$$

$$= N_0 \phi \cos l \, \zeta \left( \frac{\eta}{\eta_\lambda} \right) \Big/ \zeta \left( \frac{1}{\eta_\lambda} \right) \tag{9}$$

where $\zeta \left( \frac{\eta}{\eta_\lambda} \right) = 1 - \left( 1 - \frac{\eta}{\eta_\lambda} \right)^{1/2}$ and use has been made of (5).

An approximation can be obtained by noting

$$\zeta(\delta) \approx \tfrac{1}{2}\delta + \tfrac{1}{8}\delta^2.$$

Hence for $\eta_\lambda \gg 1$ and $\eta_\lambda \gg \eta$:

$$N = N_0 \eta\phi \cos l. \tag{10}$$

This approximation is good to within about 10 per cent provided $\eta_\lambda/\eta > 2\frac{1}{2}$ or $\sec^6\lambda/\sec^6 l > 2$ or $R > 1{\cdot}25$.

The distributions derived have been expressed in terms of $N_0$ since this quantity (as a function of $R_0$ or $\lambda$) is the one most readily found from whistler and pulsation studies. However, the $N_0(\lambda)$ distribution is derivable if we assume a distribution for $N_\lambda$, the electron density in the upper ionosphere.
Suppose

$$N_\lambda = N_E \cos \lambda.$$

Substituting in (7)

$$N_0 = \frac{N_E}{\phi_\lambda}\,\zeta\left(\frac{1}{\eta_\lambda}\right)$$

$$\approx \frac{N_E}{2\phi_\lambda\eta_\lambda} \text{ for } \lambda > 25^\circ$$

Or in terms of $R_0 (R_0 > 1{\cdot}25)$:

$$N_0 \approx \frac{N_E}{2R_0{}^3(4 - 3/R_0)}. \tag{11}$$

## 3. Discussion

### 3.1. *Assumptions*

The theory is valid provided (i) that the geomagnetic field is approximately dipole, (ii) that collisions and (iii) gravitational forces can be neglected, and (iv) that the electrons (and protons) are supplied mainly from near the earth's surface. We will consider each of these assumptions separately.

3.1.1. *Dipole field.* Recent data from the space probe Pioneer V (Coleman *et al.*, 1960) indicates that the field extends out to about 14 earth's radii. Vestine and Sibley (1959) find that the field lines connecting the auroral zones ($R_0 \sim 6$) are not seriously distorted by solar streams, even during auroral displays. This indicates the dipole field is valid out to about 10 radii.

3.1.2. *Collisions.* As pointed out in the Introduction, collisions in the regions of highest collision frequency, the source regions, are effectively taken into account by the theory. In an exosphere of protons and electrons only collisions of electrons with protons are important since elastic collisions between like particles involve only velocity exchange and the much heavier protons are not greatly effected by collisions with electrons. Assuming a temperature of about 10,000°K in the source regions, the thermal velocity of an electron would be around $10^8$ cm/sec. The time taken for one trip along the line of force $\lambda = 70^\circ$ (say) would be about 100 sec. The effective average density $\bar{N}$ ($\bar{N}s \equiv \int N ds$) along this line is about 80 cm$^{-3}$ so that at this temperature the effective average collision frequency will be about $2 \times 10^{-4}$ sec$^{-1}$. Thus the probability of an electron suffering a collision during a single trip would be about 1 in 50. Collisions would involve pitch changes and so



VOL. 20 1961

change the theoretical pitch distribution slightly. In the equilibrium condition this effect could be allowed for by a weak source in the equatorial plane giving rise to a small and essentially constant additive term (equation 3).

3.1.3. *Gravitational forces.* These can be neglected if the maximum gravitational potential energy is much less than the kinetic energy of the particles. This condition is met for temperatures in the source regions of ~2°K for electrons and ~4000°K for protons. CHAPMAN (1959) suggests that the temperature of the exosphere

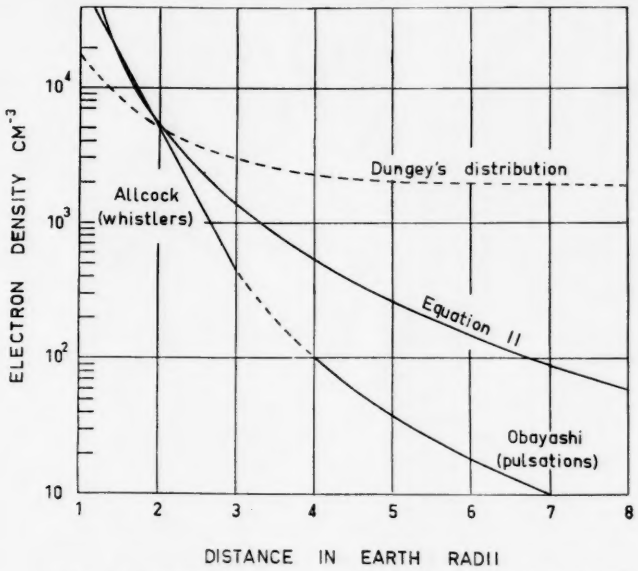


Fig. 1. Electron density along a radius in the equatorial plane.

ranges from 30,000°K at heights around 1000 km to about 200,000°K where the exosphere merges into interplanetary space. Thus it is likely that the condition is met for both electrons and protons.

3.1.4. *Source points.* JOHNSON (1959) has shown that the electrons etc. making up the exosphere must be of telluric (terrestrial) rather than solar origin. Hence the source will be the ionosphere where appreciable ionization is taking place. The theory developed above strictly requires that the "surface of the earth" be the ionosphere and so the "radius of the earth" be the radius of the ionosphere. This will effect slightly the units of $R$ and $R_0$ and the definition of $\lambda$.

### 3.2. *Computed distributions*

The electron density in the equatorial plane ($N_0$) as a function of radial distance ($R_0$) was computed from (11) for $N_E = 2 \times 10^5$ cm$^{-3}$. This value of $N_E$, the average electron density of the source region above the equator, is consistent with experimental measurements of the noon maximum density of the ionosphere ($N_{max}$) at the equator (CROOM *et al.*, 1960) of ~$20 \times 10^5$ cm$^{-3}$ and values $N/N_{max}$ ~0·2 at ~800 km (presumably the source region) found by satellite experiments (GARRIOTT, 1960). This is shown in Fig. 1 together with DUNGEY'S (1954) theoretical

distribution $N \propto \exp(2{\cdot}5/R)$ scaled to give agreement at 2 earth's radii and the experimental distribution of ALLCOCK (1959) from whistler studies for the range 1·2–3 earth's radii, and by OBAYASHI (1958) from pulsation studies for the range 4–7 earth's radii. When compared with the experimental curve we find a discrepancy which is small for $R_0 < 2$, nearly a factor of ten for $R_0 \sim 7$, but decreasing at greater distances [using OBAYASHI's extrapolation $N = \exp(17{\cdot}5/R_0)$].

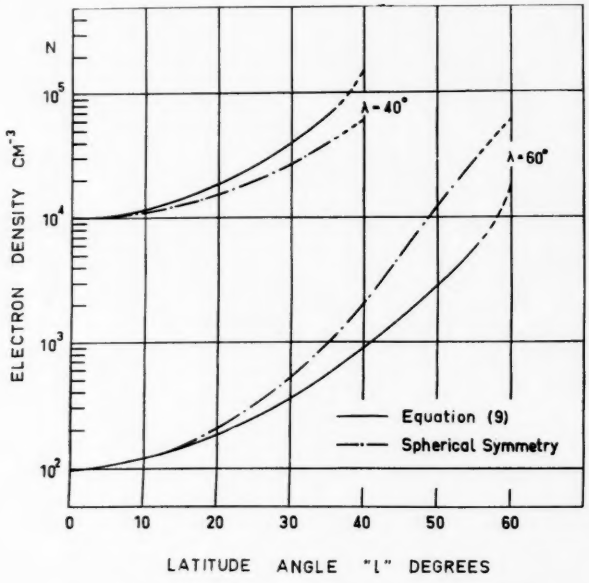


Fig. 2. Electron density along a line of force.

The distributions along lines of force $\lambda = 40°$ and $\lambda = 60°$ were computed from (9) using values of $N_0$ of 10,000 cm$^{-3}$ and 100 cm$^{-3}$ respectively taken from the whistler–pulsation experimental curve. These are shown in Fig. 2 together with distributions computed from the experimental curve on the assumption that the electron density is distributed with spherical symmetry about the earth. The dashed portions of the curves correspond to extrapolations into regions within 1000 km. The fact that the discrepancies between the two types of distribution change sign between $\lambda = 40°$ and $\lambda = 60°$ suggest that the spherical symmetry distribution may be a reasonable approximation for longitudinal propagation studies within this range of $\lambda$. However, for $\lambda < 60°$, $\phi \cos l \approx 1$ and so provided also $R > 1{\cdot}25$, we have from (10) $N \approx \eta N_0$ or $N \propto H$. Hence in this region the "whistler" mode (longitudinal extraordinary mode at frequencies much less than the gyro frequency) refractive index, being proportional to $N/H$, is approximately constant.

Values of electron density ($N$) were computed for a large number of points $(\lambda, l)$ for lines up to $\lambda = 80°$ from values of $N_0$ taken from the experimental $N_0(R_0)$ whistler–pulsation curve. For $\lambda > 65°$ values of *effective* $N_0$ were obtained by



VOL. 20 1961

extrapolation. For the longer lines of force ($\lambda > 70$) it was assumed that the theory would remain a reasonable approximation for $R < 10$ even though the line of force equations may not hold for $R > 10$. Contours of electron density or *isopycs* are shown in a plane containing the geomagnetic axis in Fig. 3. The line of force $\lambda = 70°$ is shown for reference. Presumably the contours close over the axis poles. This suggests considerable extensions of electron density in the direction of the axis.

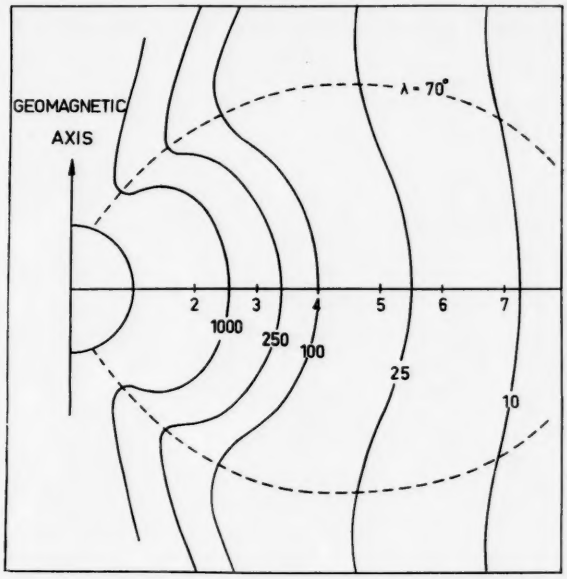


Fig. 3. Contours of electron density (isopycs) in a plane containing the geomagnetic axis.

## 4. Conclusions

The distribution of electron (and proton) density along a line of force has been derived (equation 9). At latitudes less than 60° and heights greater than about 2000 km ($R > 1{\cdot}25$) the electron density is approximately proportional to the magnetic field strength ($N/N_0 = H/H_0$). The density distribution of "aurora" or "magnetic storm" particles has also been derived (equation 3). The latter expression approximates to $N = N_0$ for latitudes less than about 60°. The electron density distribution in polar coordinates has been derived graphically (Fig. 3). At low latitudes the distribution is approximately spherically symmetrical, but at high latitudes departures from this are great, tending towards constancy along radii. This suggests considerable extensions of electron density above the geomagnetic poles.

*Acknowledgements*—The writer expresses his appreciation to Professor G. Newstead of the Department of Electrical Engineering and Dr. K. B. Fenton of the Department of Physics, University of Tasmania, and to Dr. W. G. Baker, Officer in Charge, Ionospheric Prediction Service, for helpful discussions and suggestions.

## NOTE ADDED IN PROOF

IT HAS SINCE been pointed out to the writer by Dr. J. H. PIDDINGTON and Dr. F. S. JOHNSON that satellite orbit data give temperatures of 1000–1500°K at heights around 400–700 km (HARRIS and JASTROW, 1959) so that the assumption of 10,000°K in the source regions is around seven or eight times too high. If these low temperatures applied throughout the exosphere, collision frequencies would be some twenty times greater. The electron density distribution will be largely controlled by the distribution of the much heavier protons which may not be seriously affected by electron collisions, but at these low temperatures thermal protons would have insufficient energy to neglect gravitional and geocentrifugal effects. This difficulty might be resolved by placing the source regions at a much greater height where the effective gravitional potential energy is lower and where the temperature may be much higher. The same equations derived above will describe this case provided quantities are redefined as discussed in Section 3.1.4.

On the other hand the writer has also been informed since (by Professor R. A. HELLIWELL) of a distribution of electron density out to 5 earth radii obtained from nose whistler measurements (SMITH, 1960). This distribution was found to fit a "gyrofrequency" model: $N = Kf_H\phi_\lambda^{-1}$, where $K$ is constant. Our distribution can be put in this form from (10) and (11) where

$$K = \frac{N_E\phi \cos l}{2(f_H)\lambda} \approx \frac{N_E}{2(f_H)\lambda}.$$

For $R_0 > 1{\cdot}25$ and $l$ and $\lambda < 60°$, $K$ is "constant" to within about 50 per cent. The two distributions can be closely matched throughout this region if the value of $N_E$ suggested in Section 3.2 is reduced by a factor of about three.

## REFERENCES


| | | |
|---|---|---|
| ALFVEN H. | 1950 | *Cosmical Electrodynamics*, p. 4. Clarendon Press, Oxford. |
| ALLCOCK G. MCK. | 1959 | *J. Atmosph. Terr. Phys.* **14,** 185. |
| CHAPMAN S. | 1959 | *J. Atmosph. Terr. Phys.* **15,** 43. |
| CHAPMAN S. and SUGIURA M. | 1956 | *J. Geophys. Res.* **61,** 485. |
| COLEMAN P. J., SONNETT C. P., JUDGE D. L. and SMITH E. J. | 1960 | *J. Geophys. Res.* **65,** 1856. |
| GROOM S., ROBBINS A. and THOMAS J. O. | 1960 | *Nature, Lond.* **185,** 902. |
| DUNGEY J. W. | 1954 | *The Physics of the Ionosphere*, p. 229. Physical Society London. |
| GARRIOTT O. K. | 1960 | *J. Geophys. Res.* **65,** 1139. |
| GOLD T. | 1959 | *Nature. Lond.* **183,** 355. |
| HARRIS I. and JASTROW R. | 1959 | *Planet. Space Sci.* **1,** 20. |
| JOHNSON F. S. | 1959 | *Nature, Lond.* **184,** 1787. |
| OBAYASHI T. | 1958 | *Rep. Ionos. Res. Japan* **12,** 301. |
| SMITH R. L. | 1960 | Stanford Electronics Laboratory, Techn. Rept. No. 6 |
| VESTINE E. H. and SIBLEY W. L. | 1959 | *Planet. Space Sci.* **1,** 285. |




VOL. 20 1961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 131 to 139. Pergamon Press Ltd. Printed in Northern Ireland

# Auroral radio-echoes and magnetic disturbances

C. D. WATKINS*
Jodrell Bank Experimental Station, University of Manchester

(*Received* 30 *August* 1960)

**Abstract**—It is shown that the probability of detecting auroral radio-echoes with a low sensitivity equipment at Jodrell Bank increases with the value of the magnetic $K$-index, being 100 per cent for local $K$-indices of 8 and 9. There is a close correlation between east–west moving echo regions crossing the meridian of Eskdalemuir, Scotland and magnetic disturbances recorded there. A comparison of magnetograms obtained at Lerwick and Eskdalemuir (north and south of the echo regions respectively) show that, in general, during the occurrence of echoes, the current systems giving rise to the magnetic disturbance are situated south of Lerwick and north of Eskdalemuir and thus near the echo regions.

An investigation of the location of visual auroral forms during the occurrence of echoes shows that they also are mainly situated in this region. The results confirm the suggestion of BULLOUGH *et al.* (1957) that magnetic disturbances are closely related to regions of ionization which move approximately along magnetic parallels of latitude and which can be detected with radio echo equipments.

## 1. INTRODUCTION

IN A PREVIOUS paper BULLOUGH *et al.* (1957) compared magnetic data with auroral observations obtained at Jodrell Bank ($\Phi = 56{\cdot}0°$N) with a low radio-echo equipment which used a fixed aerial directed at azimuth 292°E. Since then, auroral observations have been made with a continuously operated rotating aerial equipment of similar sensitivity (LOVELL, 1957). It has been shown by WATKINS (1960) that the auroral echoes obtained with these equipments are only obtained from a region where the line of sight intersects the local magnetic field lines perpendicularly, and that the fixed beam equipment observes only one part of this region. The rotating aerial equipment provides an almost complete coverage of the echo region and has thus enabled a fuller comparison with magnetic data to be made. The results were obtained over the period from January 1956 to September 1958. On a small number of occasions a fixed aerial system directed towards the north has been used in conjunction with the rotating aerial equipment.

## 2. GENERAL CORRELATION

The general relationship between magnetic activity and auroral echoes can be investigated by means of the magnetic $K$-index. Despite the fact that this is an index based on a 3 hourly interval, whereas echo duration can be as short as a few minutes, it is a useful quantity to use as a preliminary comparison. The distributions of values of $K$ recorded at Eskdalemuir ($\Phi = 58{\cdot}5°$N) and the planetary index $K_p$ have been obtained for the periods when echoes were being detected. The results are shown in Fig. 1. It is immediately seen that no echoes have been detected during quiet conditions i.e. $K \leqslant 2$ and that the large majority have been obtained during moderate and severe disturbances. The probability of detecting echoes has been found to be dependent on the degree of magnetic disturbance. Table 1 shows the distribution of occurrence, over the period considered, of values of $K$ at Eskdalemuir and $K_p \geqslant 6$ together with the number of

* Present address: Royal Radar Establishment, Malvern, Worcs.

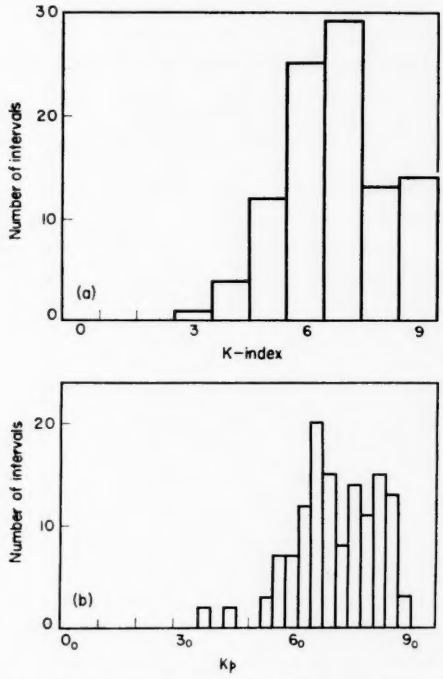


Fig. 1. (a) The distribution of Eskdalemuir $K$-indices during auroral echo activity.
(b) The distribution of planetary $K$-indices, $K_p$, during auroral echo activity.

Table 1

| Eskdalemuir $K$-index | Total No. of intervals | Intervals with echoes | Fraction with echoes (%) | Planetary $K$-index $K_p$ | Total No. of intervals | Intervals with echoes | Fraction with echoes (%) |
|---|---|---|---|---|---|---|---|
| 6 | 92 | 25 | 27 | $6_0$, $6+$, $7-$ | 184 | 39 | 21 |
| 7 | 45 | 29 | 64 | | | | |
| 8 | 13 | 13 | 100 | $7_0$, $7+$, $8-$ | 86 | 37 | 43 |
| 9 | 14 | 14 | 100 | | | | |
| | | | | $8_0$, $8+$, $9-$ | 49 | 39 | 80 |
| | | | | $9_0$ | 3 | 3 | 100 |



VOL. 20 1961

occasions when echoes were detected. The chance of detecting echoes increases markedly with the value of $K_p$ and even more sharply with the value of $K$ at Eskdalemuir. This suggests that the occurrence of echoes is more closely related to the magnetic disturbance index as recorded at Eskdalemuir than to the mean planetary index. This is as one would expect since Eskdalemuir is located just south of the region from which echoes are obtained.

The distribution of $K$ and $K_p$ in Fig. 1 have their maximum values at about 7. The magnetic $K$-index is such that over a long period the number of times a particular value occurs depends on the value considered. Thus the lower values of $K$ are in the majority with a decrease in number for increasing $K$. Since the probability of detecting echoes increases with the value of $K$ the resultant of the two variations is the maximum shown in Fig. 1.

### 3. Detailed Correlations

It was shown by Bullough *et al.* (1957) that the sign of disturbance in the $H$ geomagnetic component recorded at Eskdalemuir was always closely related to the line of sight motion of auroral echo regions as observed with the fixed aerial equipment. Furthermore, similar patterns in the variations of $H$ and echo characteristics were found. On the assumption that the line of sight motions arose from actual east–west movements of the echo regions, it was shown that the results were compatible with the statement that the echo regions are the peaks in the current systems which give rise to the magnetic disturbances, with the delay between the correlated features arising from the east–west separation of the echo regions and Eskdalemuir (about 500 km).

Because the aerial system was fixed, it was necessary to make an assumption concerning the true direction of motion of the echo regions and the observations did not therefore exclude the possibility that the motions were in fact north–south. Since it has been shown by Vegard (1912) that visual auroral forms have a pronounced diurnal north–south movement it is necessary for this point to be clarified. This has been done by observations with the rotating aerial equipment. The echo geometry at Jodrell Bank is such that echoes can only be detected from an arc which is comparatively thin in a north–south direction but covers about 1500 km of longitude (Watkins, 1960). The extremities of this arc occur at a range of about 1000 km and at azimuths of about 45°E and 65°W of north. The minimum range of about 350 km occurs towards magnetic north. On a number of occasions, echoes have been first detected from the extreme end of the arc and then gradually observed to move to shorter ranges before moving to the long ranges of the other limit of the arc. This end to end movement is very difficult to explain on the basis of a true north–south motion, whereas it is an immediate result of any east–west motion. Further evidence supporting the east–west hypothesis has been given by Unwin (1959). The origin assumption by Bullough *et al.* was therefore reasonable.

The magnetic observatory of Eskdalemuir is approximately magnetic north of Jodrell Bank. Its magnetic records are therefore most sensitive to east–west current systems due north of itself and Jodrell Bank. From the results obtained with the fixed aerial equipment, summarized above, it would be expected that

the crossing of the Jodrell Bank magnetic meridian by an echo region would correspond to a disturbance in the magnetic field at Eskdalemuir. This has been found to be the case and an example based on results obtained during the aurora of 25 February 1956 is shown in Fig. 2. During the observations, the echo regions exhibited an east to west movement and the times at which the regions were due north of Jodrell Bank have been plotted together with the $H$ geomagnetic component recorded at Eskdalemuir and Lovö (see later). It can be seen that the times at which echo regions were north of Jodrell Bank correspond closely to the times of maxima in the $H$-component. In particular, even some of the secondary

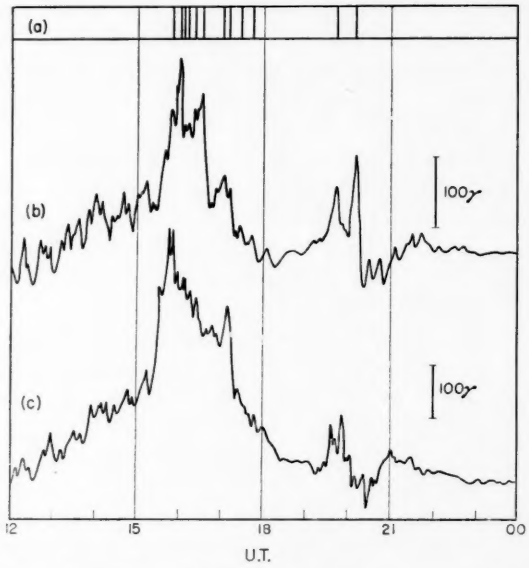


Fig. 2. (a) The times of occurrence of auroral echo regions due north of Jodrell Bank on 25 February 1956 together with the variations of the $H$-component of the geomagnetic field recorded at (b) Eskdalemuir and (c) Lovö.

maxima seem to be related to the echo regions. This correlation between localized echo regions and magnetic disturbances is also illustrated by observations made with the northerly directed aerial during the strong aurorae of 4/5 September 1958 and 8/9 July 1958. Fig. 3 shows, for both occasions, the times of occurrence of echoes detected with this aerial together with the variation of the Eskdalemuir $H$-component. It can be seen that before 2000 hours U.T. there is good agreement between echoes and maxima and after 2000 hours U.T. between echoes and minima, the maxima and minima corresponding to east to west and west to east movements of the echo regions respectively. If one now considers two locations on a similar latitude but some distance apart then, because of the relationship between echo regions and magnetic disturbances, the following features should be apparent when the magnetic variations at these locations are compared. The magnetic variations at both places should be very similar with individual features on the records having a relative time displacement depending on the separation of the locations, the sign of the disturbance and the magnitude of the echo region velocity.



VOL.
20
1961

An attempt to measure such a time displacement has been made for a number of occasions during echo activity by comparing magnetograms recorded at Eskdalemuir and Lovö, Sweden. The Lovö observatory was the closest one with a geomagnetic latitude similar to that of Eskdalemuir. It was found that the general

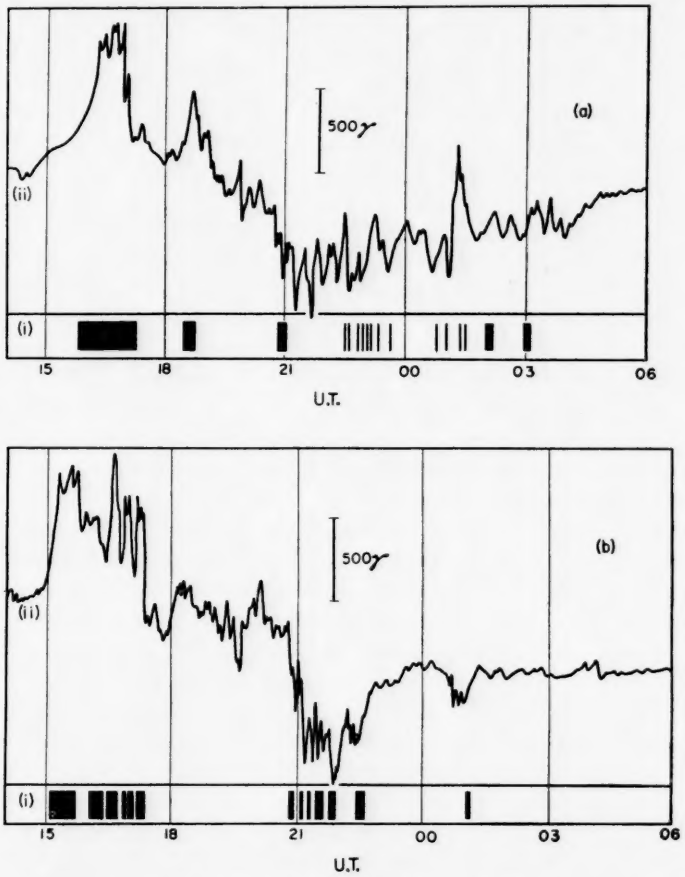


**Fig. 3.** (i) The occurrence of echoes on a northerly directed aerial together with (ii) the variations of the $H$-component of the geomagnetic field at Eskdalemuir. (a) 4/5 September 1958. (b) 8/9 July 1958.

pattern of disturbance lasting for several hours was similar at both locations but the shorter period variations lasting under an hour were generally dissimilar and could not be related. On one occasion, however, it was possible that two features were related and this occasion is illustrated in Fig. 2. The features are the two maxima occurring near 2000 hours U.T. The time differences between the respective maxima are about 10 min and 15 min and since the east–west separation of Eskdalemuir and Lovö is about 1400 km, these time differences would correspond to westward velocities of about 2000 m/sec. Measurements of the movements of echo regions at the same times indicated westward velocities of just over 1000

m/sec. The direction of motion is thus in agreement but the magnitudes of the velocities are rather different. An echo region moving with an east–west velocity of 1000 m/sec would take about 20 min to travel the distance between the two observatories. This is near the upper limit of the length of time for which echo regions have been recognized and more often the life time of such a region is less than this. Thus since echo regions related to disturbances at one observatory will probably have decayed before reaching the other it is not surprising that, in general, there is a lack of correlation between the short term features on the magnetograms at both locations. It would be expected that the correlation would be quite reasonable if the two locations were separated by a distance of about 500 km in an east–west direction.

### 4. The Latitudinal Extent of Auroral Ionization

The confinement of auroral echoes to the region of specular reflection precludes any direct evidence regarding the latitudinal extent of auroral ionization. At any location a comparison of the changes in the horizontal ($H$) and vertical ($V$) geomagnetic components indicates whether the associated current system is north or south of the position. At first sight it seems that observations from two stations situated on the same meridian can define the height and latitude of a current system but the presence of induced earth currents prevent this. The sign of the $H$-component depends only on the direction of flow of the current system and is independent of its latitude. The sign of the $V$-component, on the other hand, depends on whether the current flow is north or south of the location as well as on the direction of flow. As mentioned previously, Eskdalemuir is located just to the south of the echo region. A second magnetic observatory is located at Lerwick ($\Phi = 62{\cdot}5°$N) i.e. about 2–3° of latitude north of the echo regions and a few degrees of geomagnetic latitude east of Eskdalemuir. The magnetic records obtained at both stations have been compared for fifty days during which auroral echoes were detected at Jodrell Bank. It was found that in general the overall features of $H$ recorded at both locations were reasonably similar when one takes into account the higher degree of disturbance at Lerwick. During the quieter and less disturbed periods, the $V$-components at both locations also had reasonable agreement. During the periods when echoes were present, however, the $V$-components behaved rather differently at the two locations, with positive changes at Eskdalemuir corresponding to negative changes at Lerwick. Similarly, negative changes at Eskdalemuir were accompanied by positive changes at Lerwick. An example of this behaviour during a moderate magnetic storm is given in Fig. 4 and during a severe storm in Fig. 5. It has been shown by Bullough *et al.* (1957) that the changes in $V$ at Eskdalemuir are generally positive during the afternoon and evening hours but negative in the night and morning hours, the change-over occurring near 2100 hours U.T. In Fig. 4 it can be seen that the Eskdalemuir disturbances at about 2230 and 0130 hours U.T. are of the expected sign. The corresponding Lerwick disturbances are of opposite sign thus indicating that the responsible current systems are south of Lerwick but north of Eskdalemuir. The occurrence of echoes is observed to coincide with the times of maximum disturbance.

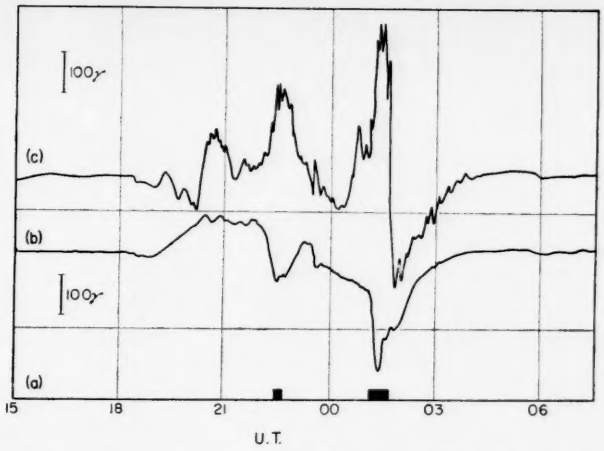


Fig. 4. (a) The times of occurrence of auroral echoes on 6/7 November 1957 together with the variations of the $V$-component of the geomagnetic field at (b) Eskdalemuir and (c) Lerwick.

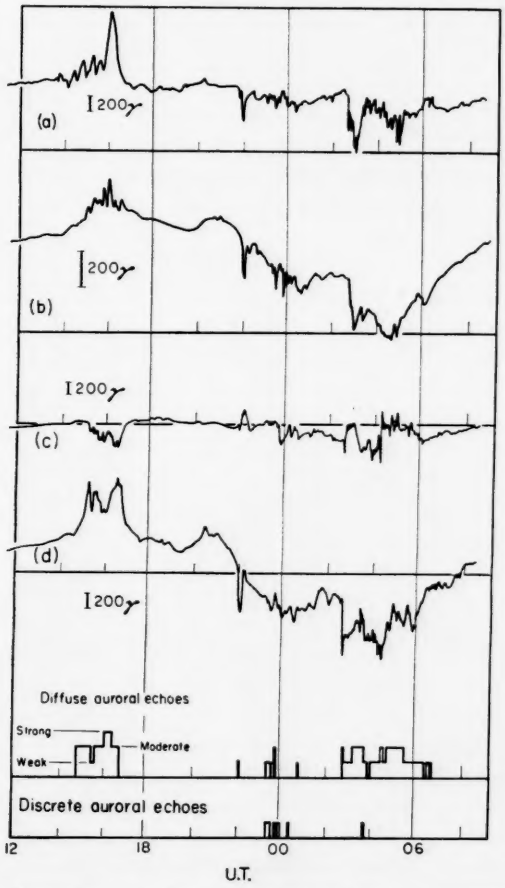


Fig. 5. The times of occurrence of auroral echoes on 22/23 September 1957 together with the variations of the $V$- and $H$-components of the geomagnetic field at Lerwick and Eskdalemuir. (a) Eskdalemuir $H$-component. (b) Eskdalemuir $V$-component. (c) Lerwick $V$-component. (d) Lerwick $H$-component.

Fig. 5 compares the $H$- and $V$- components at Eskdalemuir and Lerwick with the occurrence of echoes during very disturbed conditions. Similar general features are apparent in the $H$-components at Eskdalemuir and Lerwick and in the $V$-component at Eskdalemuir. The Lerwick $V$-component, however, shows disturbances of opposite sign during the occurrence of echoes, a particularly striking occasion being at about 2200 hours U.T. Fig. 5 also shows evidence of current systems first appearing to the north of Lerwick and then gradually moving equatorwards until they are to the south. At about 0230 hours U.T. the Lerwick $V$-component showed a negative change which suddenly reversed in sign as if a current system had passed overhead. The large negative change in the $H$-component at this time supports this suggestion, since the maximum disturbance in this component is produced by a current system exactly overhead. It is also of interest to note that a series of auroral echoes was first observed at this time.

The Eskdalemuir and Lerwick $V$-components have been studied in detail for fifty days during which auroral echoes were detected. For these days the integrated duration of auroral echoes was 6337 min and the Lerwick $V$-component was reversed, compared with the Eskdalemuir component, for a total of 9637 min. The total time when echoes and reversals occurred together was 4680 min. Thus echoes occurred during about half the time that reversals were present, and reversals occurred during about three-quarters of the time that echoes were detected. Since the analysis was restricted to days on which echoes were observed, it is possible that reversals occurring on other days have been excluded and that the former figure of about a half is too high. It is doubtful, however, whether a complete analysis of every day over the period considered would significantly alter this value. When one remembers that the echoes can only be detected from a very restricted region, the above figures show that there is quite a close correlation between current systems south of Lerwick and the detection of auroral echoes at Jodrell Bank.

Similar relationships between magnetic disturbances and visual aurorae have been investigated by MEEK (1953) and HEPPNER (1954). Both workers found that there was a close correspondence between magnetic disturbances and the presence and positions of visual forms. In particular, HEPPNER showed that the changes in the vertical component reversed in sign as the aurorae moved overhead. A similar feature has been reported by STAGG and PATON (1939).

The latitude variation of visual forms over the British Isles has been obtained for the occasions when auroral echoes were detected at Jodrell Bank after dark. For each 10 min interval with echoes, the lowest latitude with aurora in the zenith and lowest latitude from which aurora was observed were noted. The resulting variations are shown in Fig. 6. It can be seen that the maximum of the zenith curve (a) is south of Lerwick and thus coincides quite well with the positions of the previously discussed current systems during the occurrence of echoes. Fig. 6(b) is consistent with this but its resemblance to Fig. 6(c) considerably reduces its significance.

It seems, therefore, that visual aurorae and auroral radio echoes have a similar general relationship with magnetic disturbances.



VOL. 20 1961

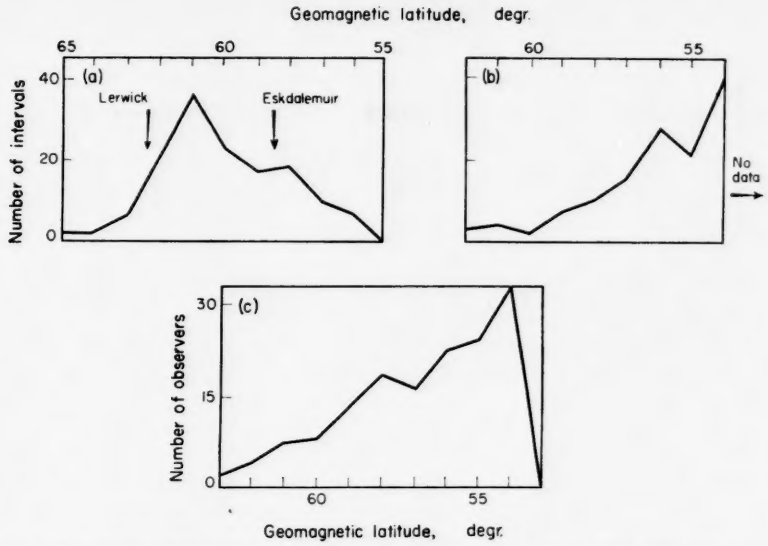


Fig. 6. The latitude distributions of: (a) Lowest latitude of aurora in zenith; (b) Lowest latitude from which aurora observed; (c) Approximate distribution of visual auroral observers.

## Conclusions

The results presented above confirm the idea put forward by Bullough *et al.* (1957) that magnetic disturbances are associated with regions of comparatively high ionization which move in an east–west direction and which are detectable as auroral radio echoes. The north–south extent of such ionization has been shown to be less than about 3° of latitude, whereas the east–west extent can vary considerably.

*Acknowledgements*—The author is indebted to the Superintendent, Meteorological Office, Edinburgh for the loan of the Lerwick and Eskdalemuir magnetograms and to Dr. N. Ambolt and Dr. F. Eleman of the Geomagnetic Section, Swedish Board of Shipping and Navigation, Stockholm for copies of the Lovö magnetograms. The author also wishes to thank Mr. James Paton and Mr. Bennet McInnes of Edinburgh University for readily making available the observations made by their wide network of auroral observers. The work formed part of the programme of the British National Committee for the International Geophysical Year.

## References


| | | |
|---|---|---|
| Bullough K., Davidson T. W., Kaiser T. R. and Watkins C. D. | 1957 | *J. Atmosph. Terr. Phys.* **11,** 237. |
| Heppner J. P. | 1954 | *J. Geophys. Res.* **59,** 329. |
| Lovell A. C. B. | 1957 | *Annals of the I.G.Y.*, Vol. 3, Pt. IV, p. 342. Pergamon Press, London. |
| Meek J. H. | 1953 | *J. Geophys. Res.* **58,** 445. |
| Stagg J. M. and Paton J. | 1939 | *Nature, Lond.* **143,** 941. |
| Unwin R. S. | 1959 | *Nature, Lond.* **183,** 1044. |
| Vegard L. | 1912 | *Phil. Mag.* **23,** 211. |
| Watkins C. D. | 1960 | *J. Atmosph. Terr. Phys.* **19,** 1. |

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 140 to 148. Pergamon Press Ltd. Printed in Northern Ireland

# Temporal variations of auroral radio-echo activity in sub-auroral latitudes

C. D. WATKINS*
Jodrell Bank Experimental Station, University of Manchester

(*Received* 30 *August* 1960)

**Abstract**—Low power radio-echo equipments have been continuously operated for several years recording observations of auroral ionization. The results have been examined for the following temporal variations of activity: diurnal, 27 day, seasonal and 11 year. The time delay between the occurrence of class 3 solar flares and subsequent echo activity has also been investigated. It is suggested that the bimodal diurnal distribution possibly arises from the condition of specular reflection which controls the detection of echoes and from the known north–south motions of visual auroral forms. The general features of the other periodicities are found to be little different from the well known ones of geomagnetic disturbance.

## 1. INTRODUCTION

CONTINUOUS radio-echo observations of the aurora borealis have been carried out over a long period at Jodrell Bank ($\Phi = 56°N$) using two equipments operating on wavelengths near 4 m. A fixed beam equipment commenced operation in 1949 and a rotating aerial equipment in 1953 and observations have been continued up to the end of the International Geophysical Year. Both sets of apparatus are of comparatively low sensitivity and have been described by ASPINALL *et al.* (1951) and BULLOUGH and KAISER (1954).

The long term nature of the observations has enabled the variation in auroral echo activity to be investigated for periodicities related to solar and magnetic effects. The rotating aerial equipment allowed a more complete coverage of the auroral echo region and, in general, the results obtained with this apparatus have been used in the analysis. The classification of the echoes into diffuse and discrete types follows that of BULLOUGH and KAISER (1955).

## 2. DIURNAL VARIATION

The diurnal variations of echo activity for (a) all echoes and (b) discrete echoes are shown in Fig. 1. This shows that the variation has two maxima, that in the afternoon hours consisting almost entirely of diffuse echoes and the morning maximum being mainly composed of discrete echoes. These features confirm similar results obtained at Jodrell Bank by HAWKINS (1954) and BULLOUGH and KAISER (1955). It has been shown by BULLOUGH *et al.* (1957) that the late evening minimum occurs at about the same time that magnetic disturbances change from positive to negative values. They therefore associated diffuse echoes with positive disturbances and discrete echoes with negative disturbances. The same apparent relationship between type of echo and sign of magnetic disturbance has been found in the more recent results. There does, however, appear to be a tendency for the type of echo to depend on the presence of sunlight rather than on the sign of magnetic disturbance. In particular there were a small number of occasions during prolonged echo activity during which the character of the echoes changed

* Present address: Royal Radar Establishment, Malvern, Worcs.



VOL. 20 1961

from discrete to diffuse at about sunrise while the magnetic disturbance remained negative. A similar solar control of echo type has been reported by BULLOUGH (1958) for Terre Adélie, Antarctica and by PRESNELL *et al.* (1959) in Alaska.

It has been pointed out by KAISER and BULLOUGH (1955) that the above variation of radio-echo activity is apparently at variance with the known variation of occurrence of visual aurorae. The diurnal variation of visual auroral activity has been studied in detail by VEGARD (1912) and he has shown that the peak activity occurs everywhere at about 1 hr before local magnetic midnight. This is in fact approximately the same time as the minimum in echo activity. A comparison between the radio-echo results and visual observations made by a large network of observers during the International Geophysical Year has confirmed this point.

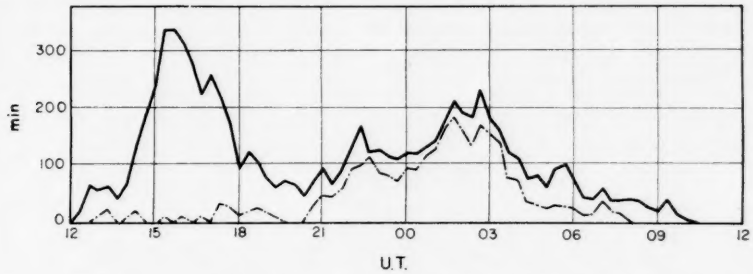


Fig. 1. The diurnal variation of auroral echo activity in minutes duration per 20 min interval. (a) all echoes, —; (b) discrete echoes only, — · —.

It has been shown by WATKINS (1960) that the auroral echoes can only be detected from a very restricted strip of sky corresponding to the region where the line of sight intersects the local geomagnetic field lines at right angles. There is thus a fundamental difference between the visual and radio-echo observations in that visual forms can be observed if they occur anywhere above the horizon whereas radio echoes are detected only if the ionization is located within the specular reflecting region.

South of the northern auroral zone, the aurora is most northerly in the early evening but thereafter gradually moves southward reaching its most southerly position near magnetic midnight (VEGARD, 1912). The reverse then takes place and the aurora returns northwards. It is the most southerly penetration that produces the maximum visual activity just before magnetic midnight and in general the more intense the aurora the further south the penetration. If one assumes that the ionization giving rise to radio-echoes exhibits a similar north–south movement then a possible explanation for the minimum in echo activity emerges. As the ionization moves south in the early part of the evening it passes through the echo region until at about magnetic midnight it is generally on the equator side of the region. Thus the probability of detecting echoes will be highest some hours before and lowest near magnetic midnight. Similarly the northward movement during the morning hours will result in a second maximum in echo activity. On this basis, however, it would be expected that weak ionization would, on a large number of occasions, penetrate only as far south as the echoing region and thus reach this position near magnetic midnight. Since the results described

above have been obtained with an equipment of only comparatively low power, it is possible that this large number of weak aurorae would have been undetected. It follows from this that if higher sensitivity equipments were used, the minimum would become less marked. Some evidence that this does indeed happen has been given by GREENHOW *et al.* (1960). They have found that using a high sensitivity apparatus at Jodrell Bank, the maximum occurrence of weak ionization of an auroral nature is at about 2200 hours U.T., i.e. near the time of minimum occurrence with the low power equipment.

While most workers at sub-auroral latitudes have observed a bimodal diurnal distribution of activity recorded by radio techniques (see for example UNWIN, 1959), GERSON (1955) has reported a variation which has a single maximum coinciding with the peak in visual activity. His results were based on the number of occasions when a widespread network of amateur observers were able to establish communication via auroral ionization. It can be seen that this situation resembles that of observing visual aurorae in that if ionization is present anywhere over a large area, some observers will be able to detect its presence. It is therefore not unexpected that this diurnal variation is similar to that of visual aurorae and the result is not incompatible with the bimodal variation obtained with a low power radio-echo equipment at a single location.

### 3. THE 27 DAY PERIODICITY

It has long been known that some types of magnetic disturbance exhibit a 27 day periodicity corresponding to the period of rotation of the sun with the recurrence sometimes lasting for many solar rotations.

The presence of a periodicity of this nature in the auroral echo activity has been examined in two ways. Firstly the incidence of echoes has been plotted in solar rotational time and is shown in Fig. 2(a). There is seen to be no part of the solar rotation during which the activity was significantly greater or less than the average activity. The preponderance of the results were obtained during the years 1956–1958 inclusive, and thus during this period there were apparently no long-enduring and well-defined regions on the sun which gave rise to auroral echo activity.

The existence of a number of active regions lasting only a small number of rotations and located at various solar longitudes would not however be revealed by Fig. 2(a). The possible existence of these is best examined by the method of super-position of epochs. The time intervals between each day with echo activity and successive days of activity have been found up to a maximum interval of 60 days. The result is shown in Fig. 2(b). It can be seen that there is no tendency whatsoever for the echo activity to recur after 27 days of after double this period. It can further be seen that there is no significant periodicity present over the time intervals considered.

It has been shown by GREAVES and NEWTON (1929) that the recurrence tendency of magnetic disturbance depends on its intensity and that for disturbances equivalent to a magnetic $K$-index greater than 7, no periodicity is observed. Auroral echoes are not often detected at Jodrell Bank when $K$ is less than 6 thus



VOL.
20
196

the absence of any recurrence tendency in echo activity is to be expected. GREENHOW *et al.* (1960) have demonstrated that equipments of very high sensitivity are capable of detecting weak ionization of an auroral nature during comparatively quiet magnetic conditions. It is quite probable therefore that the activity detected by such equipments would exhibit a 27 day periodicity and be related to the solar *M*-regions.

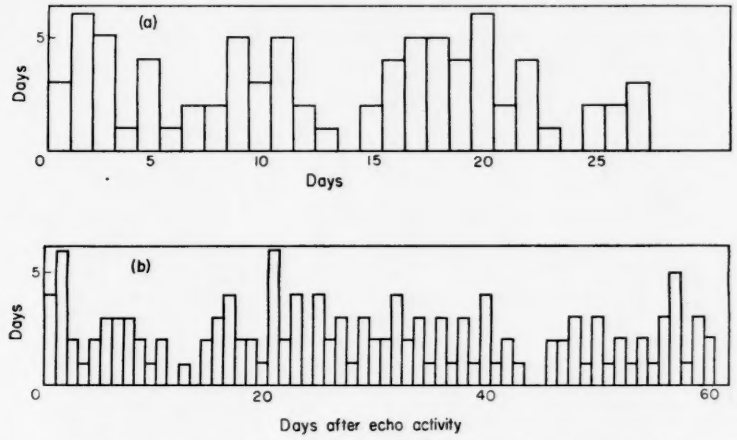


Fig. 2. (a) The incidence of auroral echoes plotted in solar rotational time. (b) The recurrence tendency of echo activity obtained by the method of super-position of epochs.

## 4. Seasonal Variation

The seasonal variation of echo activity over the period 1953–1958 is illustrated in Fig. 3 by the total number of days, hours and minutes of auroral echoes occurring in each month of the year. The seasonal variation of days of aurorae is a measure of the number of aurorae throughout the year. The variations of hours and minutes on the other hand are strongly weighted in favour of the more intense aurorae and thus are better representations of the overall auroral activity. The general features of all the curves are similar, with maxima in the spring and autumn and minima in the summer and winter. The equinoctial maxima are slightly more prominent for the curves depicting the variation of minutes of auroral echoes. Thus in addition to a larger number of aurorae occurring at the equinoxes, the average intensity of aurorae also seems to be higher at these times.

All long term visual auroral observations in sub-auroral latitudes have revealed pronounced maxima in activity near the equinoxes. The equinoctial maxima in magnetic disturbance have been known for a considerable time and the seasonal variation is apparent at all levels of activity (BARTELS, 1932). The radio-echo observations are thus in good agreement with both the magnetic and visual auroral results.

## 5. The 11 Year Cycle

The variation of echo activity over the last decade is shown in Fig. 4. The results for the years 1954–1958 have been obtained with the rotating aerial equipment and those for the years 1949–1953 with a fixed aerial equipment. The overall

sensitivities of both equipments are similar and therefore the auroral activity before and after 1953 can be compared. Fig. 4 also shows the variation of sunspot number and the number of solar flares of class 2 and greater over the period considered.

The well known 11 year period is immediately seen in the variations of sunspot number and occurrence of solar flares. The variation of echo activity is of a similar form. As one might expect, the long term agreement between sunspot number and auroral activity is quite good but there is little detailed correlation. It can be

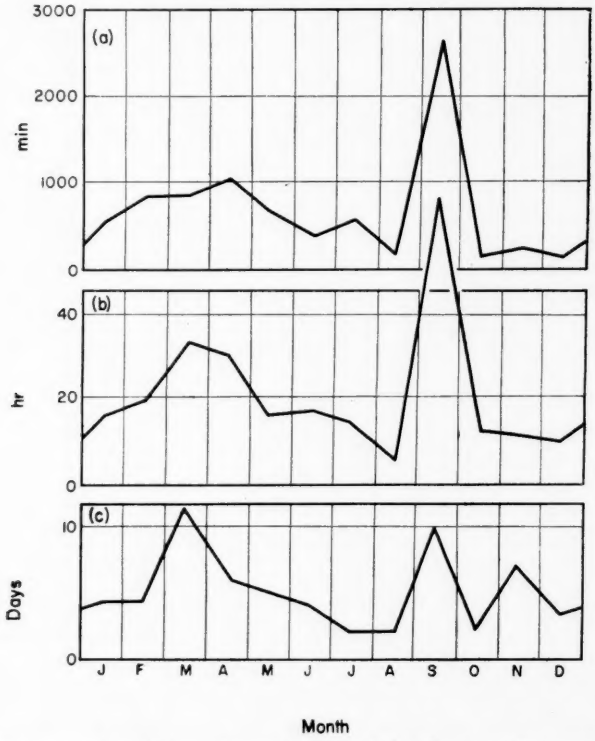


Fig. 3. The seasonal variation of auroral echo activity shown as the number of (a) minute (b) hourly (c) daily intervals per month with echoes.

seen that the sunspot minimum corresponded to a complete absence of echoes and the largest echo activity occurred near the time of the sunspot maximum in the latter part of 1957. There seems to be a slightly better correspondence between the echo activity and the number of solar flares; in particular, the maximum incidence of flares coincides in time with the highest echo activity. This suggests that the aurorae detected by radio-echo methods are of the type associated with solar flares. The relationship between the occurrence of echoes and solar flares will be discussed below.

### 6. Echo Activity Following Class 3 Solar Flares

The occurrence of echo activity has been examined in relation to class 3 solar flares during the years 1953–1958. During this period there were sixty-two such



VOL. 20 1961

flares, of which twenty-seven (44 per cent) were followed by echo activity within 5 days, and sixty-seven series of echoes, of which twenty-seven (40 per cent) were preceded by a class 3 flare within 5 days. These values are several times those which would arise from considering two unrelated series of events and this therefore suggests a relationship between the flares and echo activity. The probability of a particular flare resulting in echo activity depends on the position of the flare on the solar disc. It is found that for seventeen flares occurring within 20° of the central

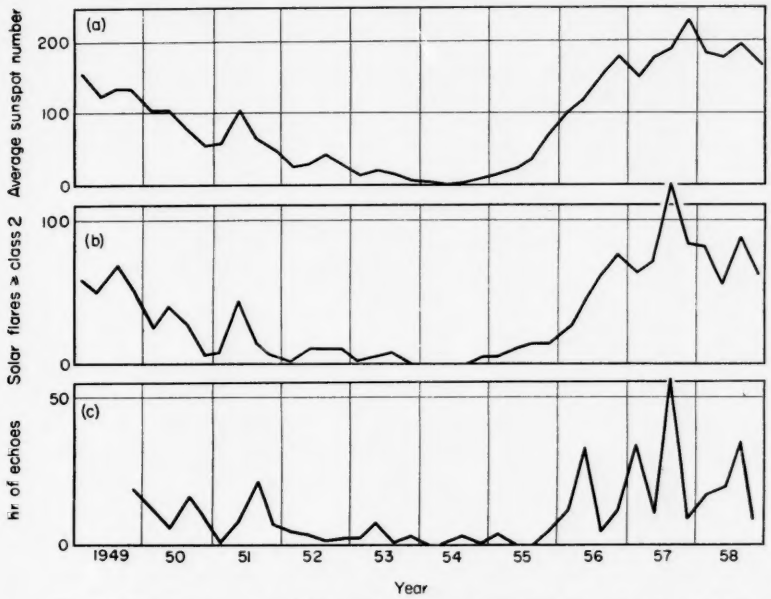


**Fig. 4.** The long term variation of (a) sunspot number (b) solar flare incidence (c) auroral echo activity.

meridian (c.m.), eleven (65 per cent) were followed by echoes; for flares between 20–40° and 40–90° from c.m. the figures were eight out of twenty-three (35 per cent) and eight out of twenty-two (36 per cent) respectively. Thus a comparatively high fraction of class 3 flares occurring near the centre of the sun's disc was followed by echo activity and this is in agreement with the idea that particles are not ejected isotropically during a flare.

The variation of echo activity in time measured from the occurrence of the flares is shown in Fig. 5. The activity is represented by the number of 2-hourly intervals with echoes. The total number of intervals with echoes was 228 and of these 127 (56 per cent) occurred within 5 days of a flare, i.e. more than would be expected from two unrelated events. Fig. 5 also shows the variation of activity for flares occurring on different parts of the solar disc. The general pattern of the variation of activity is in agreement with the accepted view that corpuscular emission from flares takes about a day or so to reach the vicinity of the earth and give rise to geophysical disturbances.

The activity seems to have two main maxima occurring at just under 2 days

and at about 3 days. This suggests that for some of the flares considered there was a delay of about 1 day before echo activity commenced and for others the delay was 2 days. This is supported by the mean variation of the magnetic index $K_p$ following the flares. Fig. 6 shows this variation for the flares within 20° of c.m. and it can be seen that the magnetic storms commence at just over 1 and just over 2 days, after

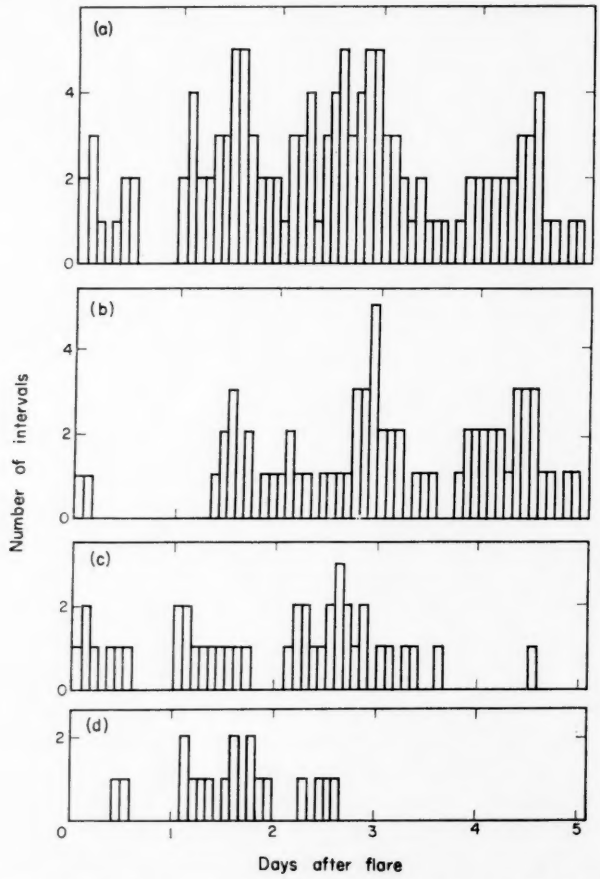


Fig. 5. The variation of auroral echo activity following class 3 solar flares. (a) all flares (b) flares within ±20° of c.m. (c) flares between ±20° to 40° from c.m. (d) flares between ±40° to 90° from c.m.

the flares. The former value has often been quoted as the time delay between a flare and subsequent geomagnetic activity but the latter figure is less common. DODSON and HEDEMAN (1958) have shown that there is a high degree of correlation between flares which exhibit sudden and intense radio bursts at 200 Mc/s frequency during the premaximum phase (termed "major early bursts") and geomagnetic storms. The average time interval between these flares and the commencement of the storms was just over 2 days. They further found that class 3 flares which did not display this type of radio emission did not give rise to magnetic disturbances.

It seems likely therefore that the sixty-two class 3 flares used to derive Fig. 5 consisted of a large fraction of "major early burst" type flares.

From above it can be seen that some 60 per cent of the series of echoes were not related to a class 3 flare. Over the period of time considered, however, there were about fifteen times as many class 2 as class 3 flares and if only a small fraction of these were capable of giving rise to geophysical disturbances then this would account

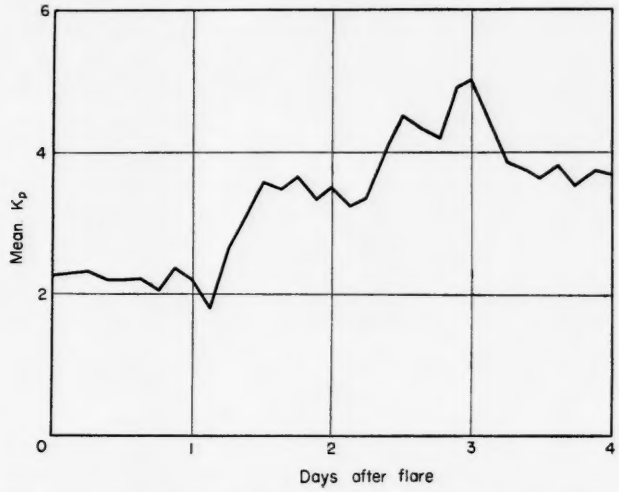


Fig. 6. The mean variation of the planetary magnetic index $K_p$ following the seventeen class 3 solar flares occurring within $\pm 20°$ of c.m.

for the above 60 per cent. It seems likely therefore that the echoes are almost always related to a solar flare occurring within the previous few days and not to a continuous particle emission from solar $M$-regions which engulfs the earth once every solar rotation. This would be in accord with the results discussed in Section 3.

## 7. Conclusions

The 27 day, seasonal and 11 year variations of auroral echo activity are little different from the well known ones associated with geomagnetic disturbances. The occurrence of echoes following solar-flare activity is also similar to that of magnetic activity.

The diurnal variation of echo activity has a bimodal distribution which possibly arises from the echo geometry and the north–south motions of the aurora.

## References


| | | |
|---|---|---|
| Aspinall A., Clegg J. A. and Hawkins G. S. | 1951 | *Phil. Mag.* **42**, 504. |
| Bartels J. | 1932 | *Terr. Magn. Atmos. Elect.* **37**, 1. |
| Bullough K. | 1958 | Report of French Antarctic Expeditions, French I.G.Y. Antarctic Sub-Committee, 47 Avenue du Maréchal Fayolle, Paris XV1. |

| | | |
|---|---|---|
| BULLOUGH K., DAVIDSON T. W., KAISER T. R. and WATKINS C. D. | 1957 | *J. Atmosph. Terr. Phys.* **11,** 237. |
| BULLOUGH K. and KAISER T. R. | 1954 | *J. Atmosph. Terr. Phys.* **5,** 189. |
| BULLOUGH K. and KAISER T. R. | 1955 | *J. Atmosph. Terr. Phys.* **6,** 198. |
| DODSON H. W. and HEDEMAN E. R. | 1958 | *J. Geophys. Res.* **63,** 77. |
| GERSON N. C. | 1955 | *Proc. Phys. Soc. Lond.* B **68,** 408. |
| GREAVES W. M. H. and NEWTON J. M. | 1929 | *Mon. Not. R. Astr. Soc.* **89,** 641. |
| GREENHOW J. S., NEUFELD E. L. and WATKINS C. D. | 1960 | *J. Atmosph. Terr. Phys.* **18,** 174. |
| HAWKINS G. S. | 1954 | *Terr. Magn. Elect. Bull.* No. 14. |
| KAISER T. R. and BULLOUGH K. | 1955 | *Ann. Géophys.* **11,** 279. |
| PRESNELL R. I., LEADABRAND R. L., PETERSON A. M., DYCE R. B., SCHLOBOHM J. C. and BERG M. R. | 1959 | *J. Geophys. Res.* **64,** 1179. |
| UNWIN R. S. | 1959 | *Ann. Géophys.* **15,** 377. |
| VEGARD L. | 1912 | *Phil. Mag.* **23,** 211. |
| WATKINS C. D. | 1960 | *J. Atmosph. Terr. Phys.* **19,** 1. |


VOL.
20
1961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 149 to 154. Pergamon Press Ltd. Printed in Northern Ireland

# Energy transport due to turbulent mixing in the atmosphere

A. J. Reynolds
Department of Aeronautics, Imperial College of Science and Technology, London

(*Received 5 September* 1960)

**Abstract**—An order-of-magnitude analysis of energy transport in a horizontally convecting turbulent atmosphere indicates that three fluxes can contribute importantly to vertical energy transfer: a heat flux due to turbulent mixing of the compressible fluid through the vertical temperature and pressure gradients; a flux of total energy produced by Archimedean forces; a work flux associated with the most important turbulent mixing stress. The relationships of these fluxes to the mean motion is discussed. It is found that the effect of the vertical pressure gradient can be represented as thermal diffusion down the potential energy gradient.

## Introduction

In a recent paper Hinze (1959) has discussed the effect of compressibility on the vertical transport of heat through the atmosphere. The transfer mechanism he analysed was turbulent mixing through the atmospheric temperature and pressure gradients. However, Yudine in the discussion following this paper called into question its results, pointing out that Archimedean forces will also contribute to the heat transfer*.

A simple analysis will be given indicating the important compressibility-dependent mechanisms of energy transfer in a turbulent atmosphere and introducing also the work flux contribution which is important in a convecting turbulent fluid.

## Notation

| | |
|---|---|
| $x$ | co-ordinate in wind direction; |
| $z$ | co-ordinate in vertical plane; |
| $t$ | time; |
| $U$ | wind speed; |
| $\Omega$ | force potential; |
| $g$ | acceleration of gravity, $\partial\Omega/\partial z$ approximately; |
| $K$ | compressibility; |
| $u, U, u'$ | velocity along $x$-axis, its time–mean value (the wind speed), and turbulent perturbation to the mean; |
| $v', w'$ | cross-wind perturbation components; |
| $p, P, p'$ | fluid pressure, its time–mean value, and the turbulent perturbation to it; |
| $\rho, R, \rho'$ | fluid density, etc.; |
| $h_0, H_0, h_0'$ | total enthalpy of the fluid, etc.; |
| $h, H, h'$ | enthalpy of the fluid, etc. |

* *Note added in proof:* Professor H. A. Panofsky has pointed out to the author that the Archimedean effect discussed here is not that intended by Yudine; the latter (a much larger contribution) makes its influence felt in the term $\overline{w'h'}$.

## The Energy Equation

We shall assume that molecular transport properties do not contribute large-scale energy flows that are significant in comparison with those set up by turbulent mixing. This assumption takes analytic form in setting the coefficients of viscosity and thermal conductivity equal to zero in the energy equation. Molecular diffusion may, indeed, have an important secondary effect by influencing lumps of fluid moving through the turbulent field. We shall suggest how it may be taken into account at a later stage in the analysis.

Neglecting molecular transport properties and assuming that motion takes place in an irrotational force field, we can write the energy equation for a moving fluid in the form

$$\rho \frac{D}{Dt}(H_0 + \Omega) = \frac{\partial p}{\partial t} + \rho \frac{\partial \Omega}{\partial t}$$

where $h_0 = h + \frac{1}{2}\underline{u}^2$ is the total enthalpy of the fluid, $\underline{u}$ being the vector velocity, and

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \underline{u} \cdot \nabla$$

is the mobile operator.

Taking mean values with respect to time and requiring that the mean motion and the force field be steady, we get

$$\nabla \cdot [\overline{\rho(h_0 + \Omega)\underline{u}}] = 0$$

where the bar above indicates that a time mean is to be taken.

We consider turbulent mixing in a moving atmosphere, dealing for simplicity with a flow essentially two-dimensional as well as steady. We let the wind speed be $U$; it may vary with altitude. The $x$-axis is chosen in the wind direction and the $z$-axis upwards in the vertical plane. Taking account of horizontal symmetry, we obtain the energy equation

$$\frac{\partial}{\partial x}[\overline{\rho(h_0 + \Omega)u}] + \frac{\partial}{\partial z}[\overline{\rho(h_0 + \Omega)w'}] = 0. \tag{1}$$

We have

$$\rho = R + \rho', \; u = U + u', \; h_0 = H_0 + h_0'$$

$$H_0 = H + \tfrac{1}{2}[U^2 + \overline{u'^2} + \overline{v'^2} + \overline{w'^2}]$$

$$h_0' = h' + u'U$$

the mean values of the cross-wind components being zero.

## Order-of-Magnitude Analysis

We shall derive approximate forms for the energy equation using an order-of-magnitude analysis.

Units are chosen such that $U \sim 0(1)$, $R \sim 0(1)$; and we allow the velocity and density changes in the field to be as large as 0(1). The magnitude of the energy gradients will not be specified; it is suggested by the analysis.



VOL. 20 1961

The turbulence is specified by $u',v',w' \sim 0(\beta)$; $\rho' \sim 0(\eta)$; $h' \sim 0(\alpha)$; and we take also $\partial/\partial x \sim 0(1)$, $\partial/\partial z \sim 0(\delta)$. It is assumed, of course, that $\alpha$, $\beta$, $\delta$, $\eta$ are all small compared to unity so that terms of higher orders in them can be neglected.

The assumption of small-scale turbulence rules out the application of the results to the very large-scale mixing to be expected in an atmosphere which is not stably stratified.

The relationship between the gradients is crucial to the simplifying analysis. It is not unlike the boundary layer approximation and does, in fact, imply that the $z$-axis be nearly vertical, the wind very nearly horizontal. The wide applicability of the assumption is a result of the fact that the earth's atmosphere is a very thin film.

Introducing these orders into the energy equation (1), we find that

$$\frac{\partial}{\partial x}[(H_0 + \Omega)RU] + \frac{\partial}{\partial x}[\overset{0(\alpha\beta)}{\overbrace{R\overline{w'h'}}} + (H_0 + \Omega)\overset{0(\beta\eta)}{\overbrace{\overline{\rho'w'}}} + \overset{0(\beta^2)}{\overbrace{RU\overline{u'w'}}}] = 0(\delta\alpha\beta,\ \delta\beta^2,\ \delta\alpha\eta,\ \delta\beta\eta). \tag{2}$$

Using the mean-value continuity equation, we can cast this result into another form without altering the orders of the error terms. Thus

$$RU\frac{\partial}{\partial x}(H_0 + \Omega) = R\frac{D}{Dt}(H_0 + \Omega) = -\frac{\partial}{\partial z}(R\overline{w'h'}) - \overline{\rho'w'}\frac{\partial}{\partial z}(H_0 + \Omega) - \frac{\partial}{\partial z}(RU\overline{u'w'}). \tag{3}$$

The form (2) is a relation among the several important energy fluxes through a fixed volume of the fluid, while the alternative (3) relates the energy input to the convecting fluid to the vertical fluxes.

## The Important Energy Fluxes

(a) The heat flux $R\overline{w'h'}$ is the effect of vertical mixing through the atmospheric temperature and pressure gradients which was discussed by Hinze. At first sight it appears to be simply the turbulent analogue of thermal conduction due to molecular diffusion and thus might be expected to have the same qualitative effect, to even out temperature gradients. However, work is done by or on the moving lumps of the compressible fluid as they expand or are compressed to match the local pressure. Thus the overall effect of this term is to drive the temperature distribution towards that defined by an adiabatic law applied to the pressure gradient.

(b) The flux of total energy $(H_0 + \Omega)\overline{\rho'w'}$ is due to Archimedean forces; this is the contribution suggested by Yudine. The correlation $\overline{\rho'w'}$ may be expected to be negative since the lighter particles will tend upwards under the action of buoyancy forces.

(c) The work flux $UR\overline{u'w'}$ is associated with the most important of the turbulent mixing stresses.

The relative importance of these contributions will differ widely in various

situations. Note that the Archimedean effect is wholly dependent on the compressibility of the turbulent fluid. On the other hand, the work flux is independent of compressibility. The heat transfer term exists (although in a modified form) even in an incompressible fluid.

For the case in which the wind speed is zero ($U = 0$) the steady-state analysis just given seems inappropriate. However, it does indicate what will be the determining factors in the energy flux for that case. The work flux ceases to contribute to the energy flow. The Archimedean effect and mixing through the vertical gradients remain as the efficient mechanisms of energy transport.

## Effects of Mean Gradients

Having isolated the important energy transport mechanisms, our next task is to relate them to the mean properties of the flow—the pressure, velocity and energy gradients.

(a) *The heat transfer coefficient*

The relation of the thermal mixing correlation $\overline{w'h'}$ to the temperature and pressure gradients has been given by Hinze, among others. Assuming the mixing fluid to be a perfect gas, Hinze (1959) obtained a result equivalent to

$$\overline{w'h'} = -\epsilon_H \left( \frac{\partial H}{\partial z} - \frac{\gamma - 1}{\gamma} \frac{H}{P} \frac{\partial P}{\partial z} \right)$$

where $\gamma$ is the ratio of specific heats and $\epsilon_H$ is a positive coefficient. The second term on the right accounts for work done by and on the rising and falling lumps.

Using the equation of state of the gas and the simple condition of vertical equilibrium (which is very nearly exact in this nearly horizontal flow), we can write alternative forms for this correlation:

$$\begin{aligned} \overline{w'h'} &= -\epsilon_H \left( \frac{\partial H}{\partial z} - \frac{1}{R} \frac{\partial P}{\partial z} \right) \\ &= -\epsilon_H \left( \frac{\partial H}{\partial z} + g \right) = -\epsilon_H \frac{\partial}{\partial z} (H + \Omega). \end{aligned} \tag{4}$$

It should be noted that the coefficient $\epsilon_H$, although primarily a representation of the turbulent activity within the fluid, can also express the molecular transport properties of the fluid. In deriving the expression for $\overline{w'h'}$, Hinze has shown that non-adiabatic lump movements can be taken into account. Thus our early neglect of molecular transport terms (on the grounds that their large-scale effects must be small) does not finally suppress the role of viscosity and conductivity. Molecular transport can have an important influence on the small-scale movements of lumps. Hence, its effects are probably best introduced in the coefficients giving the effects of lump movements through the mean flow.

(b) *The Archimedean correlation*

The correlation $\overline{\rho'w'}$ owes its existence to the net upwards motion of lighter lumps of the turbulent fluid. The upwards drift is superposed on random

movements in the turbulent field. It is this drift speed which contributes to the correlation.

It seems likely that the drift speed of a particle will be directly proportional to the bouyancy force on it. This relationship between net lump motion and driving force is suggested by analogy with other problems in which small particles are subjected to random fluctuations while moving in a steady force field. Such problems occur in electricity (electron drift superposed on thermal agitation) and in the theory of Brownian motion. The linear resistance laws of Ohm and Stokes are applicable there.

By analogy with Stokes' law as applied by EINSTEIN (1905) in the study of Brownian motions, we relate a typical density fluctuation and a typical drift speed by

$$\rho' g \propto R\epsilon_M w_d$$

where $w_d$ is the drift speed and $R\epsilon_M$ replaces the molecular viscosity, $\epsilon_M$ being the eddy viscosity. The constant of proportionality is dependent on the scale of the turbulence.

We shall go one step farther in attempting to separate the effects of turbulence from the basic features of the flow. It seems plausible that turbulent density fluctuations at a point are the result of the convection past that point of eddies in which density gradients are maintained by the centripetal pressure gradients holding the eddies together. This idea is put in analytic form by

$$p' \propto R(\overline{u'^2} + \overline{v'^2} + \overline{w'^2})$$

with $\rho'/p' \simeq K$, the compressibility of the fluid.

Then

$$\overline{\rho' w'} \propto -\rho' w_d \propto -RK^2 g$$

and, finally,

$$\overline{\rho' w'} = -\epsilon_A RK^2 g \tag{5}$$

where $\epsilon_A$ is a positive coefficient representing the kinematic aspects of the turbulence. It should be noted that $K^2 g$ will be essentially constant through most of an atmosphere like that of the earth.

(c) *The shear stress correlation*

The apparent stress $R\overline{u'w'}$ can be related to the mean velocity gradient by several hypotheses. The introduction of the eddy viscosity is the most convenient here. We write

$$\overline{u'w'} = -\epsilon_M \frac{\partial U}{\partial z} \tag{6}$$

where $\epsilon_M$ is a positive quantity expressing the ability of the turbulence to accomplish momentum transfer. However, as Prandtl's mixing length arguments show, the eddy viscosity is not independent of the mean motion.

We have attempted to express each of the important correlations as the product of a term representing the gross features of the flow and a positive coefficient expressing the turbulent activity within the flow. Although these results do not

allow us to give numerical values to the correlations—the exact nature of the eddy coefficients is still a mystery—they do allow us to determine the direction of each of the energy fluxes if a particular mean motion is specified.

Introducing the correlations as given in (4), (5) and (6) into the approximate energy equation (3), we get

$$R\frac{D}{Dt}(H_0+\Omega)=\frac{\partial}{\partial z}\left[R\left\{\epsilon_H\frac{\partial}{\partial z}(H+\Omega)+\epsilon_M\frac{\partial}{\partial z}\left(\frac{U^2}{2}\right)\right\}\right]$$
$$+\epsilon_A RK^2\frac{\partial\Omega}{\partial z}\frac{\partial}{\partial z}(H_0+\Omega).$$

The first term on the right is of the form that would be given by the straightforward introduction of coefficients to represent the diffusion of heat and kinetic energy by turbulence. The coefficients $\epsilon_M$ and $\epsilon_H$ are not equal since the mechanisms of transfer differ.

It has been possible to give a surprisingly simple view of the effect of the vertical pressure gradient on the energy transport. It is interpreted here as a flow of energy down the potential energy gradient, analogous to thermal diffusion. It should be noted that this interpretation is valid only for a fluid obeying the perfect gas law reasonably well. For example, the result cannot be applied meaningfully to an incompressible fluid, even though the compressibility does not enter explicitly.

The second term on the right, the Archimedean effect, is here represented as the interaction between the potential and total energy fields within a compressible fluid. This effect could not easily have been predicted without considering the details of the mixing.

## Discussion

The problem dealt with here is a highly idealized one. However, a similar order-of-magnitude analysis can be applied to pick out the significant effects in more complex turbulent mixing situations. The method is also enlightening when applied to Reynolds' momentum equations for complicated situations.

The energy separation in a vortical flow of a compressible fluid can be treated in this manner; this is the problem of the Ranque–Hilsch vortex tube. The physical effects found to be of interest in atmospheric mixing also give rise to the important energy transfers in the swirling flow. Once again, the pressure gradient normal to the mean streamlines plays an important part and Archimedean forces complicate the situation.

## References


Einstein A. 1905 *Ann. Phys. Lpz.* **17,** 549.

Hinze J. O. 1959 *Advances in Geophysics*, Vol. 6, p. 229. Academic Press, New York

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 155 to 166. Pergamon Press Ltd. Printed in Northern Ireland

# Random fluctuations in very low frequency signals reflected obliquely from the ionosphere

J. K. HARGREAVES
D.S.I.R. Radio Research Station, Slough*

(*Received* 13 *September* 1960)

**Abstract**—The paper describes the analysis of observations made on the amplitude fluctuations of very low frequency signals received from transmitters at Varberg, Sweden, and Annapolis, U.S.A. The observations were made at numerous receiving points simultaneously at ranges between 800 and 6000 km from the transmitters. The fluctuations in amplitude were of the order of 0·15 of the mean amplitude at 1000 km range and of the order of 1·02 at 6000 km range. From the calculated correlation between observations at two separated locations the size and movements of the ionospheric irregularities causing the fluctuations were deduced. The correlation patterns were between 50 and 120 km wide and the fading times of the order of 20 min. The fluctuations were due mainly to changes in their structure rather than to drifts. The measured changes in the phase difference between the signals received at two separated points have been compared with those deduced from the amplitude measurements and good agreement was obtained.

The nature of the changes in the ionosphere which give rise to the fluctuations have been studied with the aid of data obtained from phase and amplitude measurements at short distances from the 16 kc/s transmitter at Rugby. It is found that the ionosphere modulates the phase rather than the amplitude of the wave and that a model of the ionosphere in which the reflection coefficient is constant and the height fluctuates is in conformity with the observed fluctuations at the ground.

VOL.
20
1961

## 1. INTRODUCTION

ALL RADIO signals received after reflection at the ionosphere are subject to fading which can be attributed to the movement of irregular patches of ionization. BOOKER *et al.* (1950) explained the fading in terms of an irregular reflecting surface which modulates the incident wave in amplitude or phase, or both, and related the statistical properties of the signal variations observed on the ground to those at the reflector. All the relevant theory was reviewed by RATCLIFFE (1956) and compared with the available experimental data which had been obtained from observations at nearly vertical incidence.

The technique which has been adopted for the study of the irregularities is that of making simultaneous observations of the signal received at several geographically separated sites. The size of the irregularities and their movements can then be deduced by comparison between individual records of amplitude fading. Hitherto the investigations of v.l.f. fading (HARWOOD, 1953; BOWHILL, 1956) have been based on observations made at nearly vertical incidence and with small separations. The present work extends the study by measurements at very low frequencies over greater ranges and receiver separations.

To deduce the process in the ionosphere which originally caused the fading is more difficult. BOWHILL (1957) dealt theoretically with the two simplest cases, in which only the amplitude or the phase of the incident wave was changed, and deduced the consequences at the ground from diffraction theory. In the second part (Section 3) of the present paper are given further results of very-low-frequency

* Official communication.

measurements and it is shown that a simple model of the ionosphere is adequate for the explanation of the irregularities.

## 2. FLUCTUATIONS IN THE SIGNALS REFLECTED FROM THE IONOSPHERE

### 2.1. *Measurements of the amplitude of v.l.f. signals*

In previous measurements (HARWOOD, 1953; BOWHILL, 1956) on signals reflected at nearly vertical incidence from the lower ionosphere it was the horizontally-polarized component of the elliptically-polarized ionospheric wave which was recorded. At ranges of about 1000 km and 6000 km from the transmitter, at which the present measurements were made, the horizontal component is negligible and,

Table 1

| Transmitter | Season | Amplitude | | | | Relative phase | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Observing periods | Sites | Ranges (km) | Separations (km) | Observing periods | Sites | Ranges (km) | Separations (km) |
| Varberg | Aug. 1956 | 4 | 4 | 841–1121 | 83–284 | 2 | 2 | 841,1044 | 206 |
| $f = 17{\cdot}20$ kc/s<br>$\lambda = 17{\cdot}45$ km | Nov. 1956 | 4 | 4 | 1038–1046 | 32–245 | 4 | 2 | 1038,1044 | 245 |
| | Jan.–Feb. 1957 | 8 | 6 | 841–1392 | 16–364 | | | | |
| | Mar.–Apr. 1957 | 7 | 4 | 841–2371 | 206–304 | | | | |
| | July 1957 | 7 | 5 | 841–2112 | 206–555 | 4 | 2 | 1038,1044 | 245 |
| | Oct.–Nov. 1957 | 10 | 6 | 1039–2556 | 12–555 | | | | |
| Annapolis | Nov.–Dec. 1957 | 18 | 4 | 5815–5825 | 31–293 | | | | |
| $f = 15{\cdot}50$ kc/s<br>$\lambda = 19{\cdot}34$ km | Jan. 1958 | 20 | 4 | 5430–5590 | 170–800 | | | | |
| | Feb.–Mar. 1958 | 15 | 2 | 5760,5820 | 286 | 6 | 2 | 5760,5820 | 286 |

therefore, it was the total vertical component of the field on which the observations were made. Because the composition of the signals received at 1000 km is relatively simple, properties of the ionosphere can still be inferred from the observations. At ranges of 6000 km, however, the problem is more complex.

A large number of observations of the variations in the amplitude of signals from the transmitters at Varberg, Sweden (17·2 kc/s) and at Annapolis, U.S.A. (15·5 kc/s) have been made at ranges between 840 and 5820 km at different seasons of the year. Table 1 shows the scope of the observations with the ranges of the receiving sites and the separations between them. Details of the recording equipment and methods are given elsewhere by PRESSEY *et al.* (1960).

The amplitude records were in most cases continuous through the 24 hr but only the night periods, clear of the systematic effects associated with sunset and sunrise, have been used in the analysis. In the case of Varberg daytime fluctuations are very small and not easily separated from the diurnal trends. The Annapolis observations were confined to winter so as to obtain long periods of total night, and although the daytime fluctuations are considerable at the larger ranges, they are still much less than at night. The observations were usually made at two or more sites simultaneously.



VOL. 20 1961

### 2.2. *Magnitude and distribution of fluctuations*

The magnitude of the observed fluctuations is measured by $a/A$, where $a$ is the standard deviation of the amplitude from the mean amplitude, $A$. Fig. 1 shows the dependence of $a/A$ on range for both Varberg and Annapolis night periods. Each point is the average of all observations at that range and the standard error, estimated on the basis of the length of the record, is indicated by the vertical line. The readings used in the calculation of the standard deviations were taken at 5 min intervals throughout the observing period. The longest individual records, which are of 10 hr duration, give values of $a/A$ which spread about the mean by about

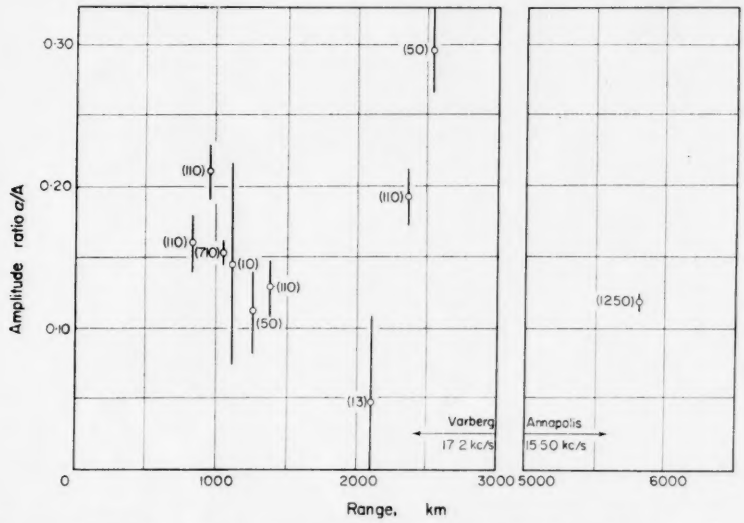


Fig. 1. Magnitude of fluctuations at night. Figures in brackets denote total number of hours of recordings.

$\pm 0{\cdot}06$. One of the most accurate figures is that for 1000–1100 km range, where a value of $a/A = 0{\cdot}153 \pm 0{\cdot}007$ was obtained from 710 hr recording. There is no significant variation with season.

The signal received at 1000 km from the Varberg transmitter is, according to the theoretical work of WAIT and MURPHY (1956) composed of ground wave, one-hop and two-hop ionospheric waves having amplitudes in the ratio $1{\cdot}0:2{\cdot}5:0{\cdot}5$, the first two being approximately in antiphase at night. Hence the value of $a/A$ for the one-hop component is about two-thirds the value obtained for the resultant signal, i.e. $0{\cdot}1 \pm 0{\cdot}004$.

The shape of the amplitude distribution curve plotted from the observations was found to be that of a displaced Gaussian which is to be expected for a randomly varying component of signal superimposed on a steady signal (see Section 2.5).

### 2.3. *Size of irregularities*

The correlation coefficient, $\rho(d)$, between the fluctuations at two receiving sites separated by a distance, $d$, is shown in Figs. 2(a)–(c). The Varberg results are given

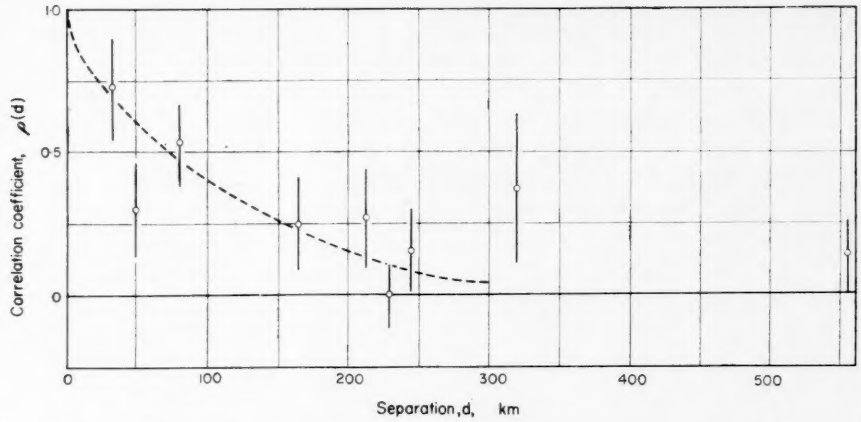


Fig. 2(a). Space correlogram—Varberg transmitter. Orientation perpendicular to propagation direction. Range 1000 km; $f = 17{\cdot}2$ kc/s; — — — fitted time correlogram.

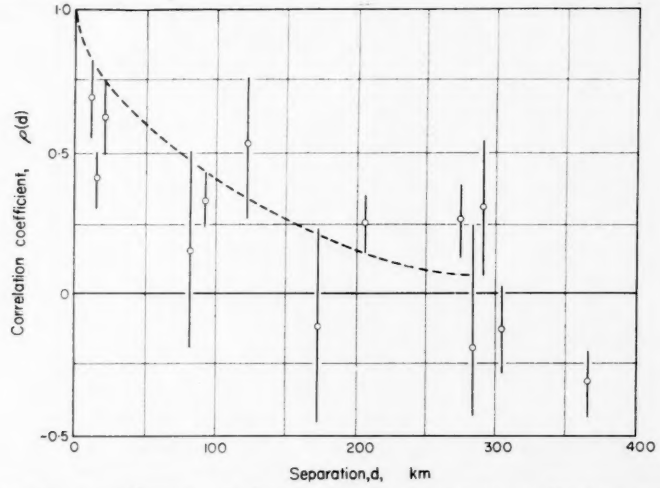


Fig. 2(b). Space correlogram—Varberg transmitter. Orientation in direction of propagation. Range 1000 km; $f = 17{\cdot}2$ kc/s; — — — fitted time correlogram.

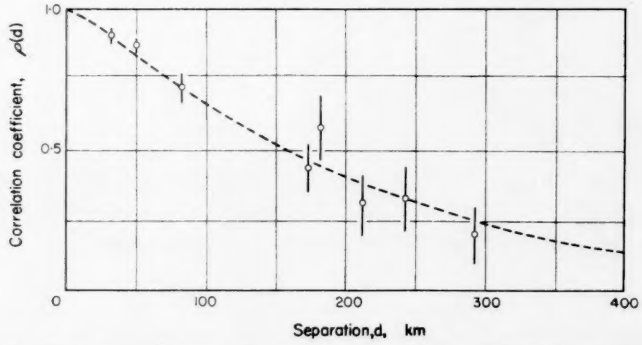


Fig. 2(c). Space correlogram—Annapolis transmitter. Orientation perpendicular to direction of propagation. Range 5800 km; $f = 15{\cdot}50$ kc/s; — — — fitted time correlogram.

in Fig. 2(a) when the sites were at a constant radius of about 1000 km from the transmitter and in Fig. 2(b) when they were approximately along the direction of propagation. Fig. 2(c) shows the results for Annapolis, the range being virtually constant at 5800 km. Each value shown is the average of several records and the standard error of the mean is shown in each case by the length of the vertical lines through the points.

The correlation distance, $d_0$, is taken to be that separation at which the correlation coefficient has fallen to $0{\cdot}61[=\sqrt{(1/e)}]$. This distance is estimated to be $46 \pm 8$ km for 17·2 kc/s at 1000 km and $115 \pm 12$ km for 15·5 kc/s at 5800 km. The distance is independent of orientation, and the large scatter precludes any investigations of a seasonal effect. BRIGGS *et al.* (1950) showed that the correlation distance at the ground is twice that at the height of the reflecting layer when the source of radiation is a point, provided only one reflection is involved. The 1000 km result, therefore, gives the size of the irregularities of the ionospheric structure observed on 17·2 kc/s as $23 \pm 4$ km. It is to be noted that in the analysis the contribution to the fading by the two-hop ionospheric wave and any fluctuations in transmitter output have been neglected.

### 2.4. *Fading*

Fig. 3 shows the average auto-correlation coefficient $\rho(t)$ as a function of time interval for the Varberg and Annapolis transmissions. The scatter of individual results is again about $\pm 0{\cdot}3$ from the mean. The records were analysed at 5 min intervals giving 120 readings in a 10 hr period. However the slow fading implies that these readings are not independent and in fact even the largest periods of observation are statistically short. An analysis of the interdependence of the readings showed that a standard error of the order of 0·4 might be expected in the derivation of the correlation coefficients. The observed scatter could be explained, therefore, by the short duration of the records and is no indication that the statistical properties of the ionosphere vary to this extent.

The fading time at which $\rho(t) = 0{\cdot}61$ is $18 \pm 4$ min at ranges near 1000 km, and $26 \pm 4$ min at 5800 km. The Varberg results shown include ranges from 840 to 1390 km but any variation of fading time with range or with season is masked by the large scatter of the observations.

It is to be noted that in each case the shape of the time and phase correlograms are similar as is shown by the fit of the time correlogram to the corresponding space correlogram in Fig. 2.

### 2.5. *Relation between amplitude and phase fluctuations in the received signals*

The concept of a random signal described by RATCLIFFE (1956), BRAMLEY (1951) and BOWHILL (1957) is extremely useful in relating the amplitude and phase properties of a fluctuating signal. The ionospheric wave is considered as consisting of two parts. One is steady. The amplitude of the other varies with a Rayleigh distribution and all phase angles relative to the steady component are equally probable. Provided that the random component is much smaller than the steady one, which is true for our cases (see Section 2.2), the standard deviation of phase in

radians, $\phi$, is equal to the standard deviation of amplitude expressed as a fraction of the mean amplitude, $a/A$.

i.e. $$\phi = a/A. \tag{1}$$

Moreover, BRAMLEY (1951) showed that the space correlograms of the phase and amplitude are identical.

All the amplitude records taken on the Varberg and Annapolis transmissions have been examined to determine whether the fluctuations are random in the above sense. From simultaneous amplitude observations at two receivers, 1 and 2,

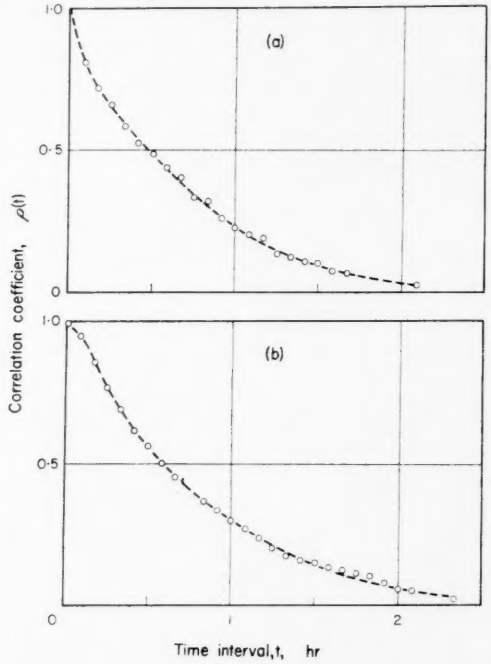


Fig. 3. Time correlograms. (a) Varberg transmitter. Range = 850–1390 km; $f$ = 17·2 kc/s; 450 hr observations. (b) Annapolis transmitter. Range = 5800 km; $f$ = 15·50 kc/s; 300 hr observation.

separated by distance $d$, the magnitude of the fluctuations $a_1/A_1$, $a_2/A_2$, and distance correlation coefficient of amplitude $\rho_a(d)$, were determined. Assuming a random signal the corresponding phase quantities can be computed,

$$\phi_1 = a_1/A_1; \quad \phi_2 = a_2/A_2; \quad \rho_\phi(d) = \rho_a(d).$$

The standard deviation of the difference of phase, $\phi_{(1-2)}$, which has also been observed, is calculated from

$$\phi_{(1-2)} = \sqrt{[\phi_1^2 + \phi_2^2 - 2\phi_1\phi_2\rho_a(d)]}. \tag{2}$$

The values of phase differences computed in this way have been compared with measured values of phase for all series of observations on both Varberg and

Annapolis signals which included simultaneous measurements of phase and amplitude (see Table 1). The comparisons, which are described in detail by PRESSEY *et al.* (1960) showed good general agreement. The daily variations of measured and calculated phase also correlate with a coefficient of 0·76 at 1000 km indicating that there is a day-to-day, besides overall, correspondence. Thus at these distances the amplitude records give a good indication of the extent and nature of the phase fluctuations.

2.6. *Motion of the irregularities*

The signal received at the ground after reflection from an irregular reflecting layer fades because the irregularities move about, or change their form, or do both. BRIGGS *et al.* (1950) defined the "Drift velocity", $V$, to represent movement as a whole, and the "Characteristic velocity", $V_c$, to represent changes of structure. Neither can be measured directly, but they are related to the "Fading velocity", $V_c'$, defined as the ratio between fading distance and fading time ($d_0/t_0$ in our terms), and "Apparent drift velocity", $V'$, equal to the separation between two observing points, $d$, divided by the time displacement, $T$, which produces maximum correlation between their respective amplitude records.

The relations are

$$VV' = (V_c')^2 \tag{3a}$$

$$(V_c')^2 = V^2 + V_c^2. \tag{3b}$$

A series of six sets of observations made simultaneously at three stations separated by less than 25 km have been analysed to determine the values of time displacement in two approximately orthogonal directions. The observations were made on the Varberg signals at a range of 1000 km in September 1957.

The apparent drift velocities, $V'$, were calculated from these measured time displacements and from the values of $d_0$ and $t_0$ already obtained at this range (see Sections 2.3 and 2.4) the fading velocities, $V_c'$, were also determined. The components of the drift velocity, $V$, in the directions of the receiver baselines were then obtained by substituting in (3a). Values for the resultant drift velocity lay between 7 and 21 m/sec with an average of 15 m/sec. The directions of the velocity vectors were southerly except in one case in which the direction was northerly. The average value for the characteristic velocity. $V_c$, obtained by substitution of $V_c' = 43$ m/sec in (3b), was 40 m/sec. The corresponding values at the reflecting layer are half these ground values. These limited data indicate that, since the characteristic velocity is greater than the drift velocity, the fading is caused more by changes of structure than by drift motion.

A similar examination of observations made in January 1957 at three stations separated by about 200 km produced drift and characteristic velocities of the same order.

The whole night period was used in these analyses so that only the average drift of the ionospheric irregularities during the night was obtained. To shorten the records further in order to obtain more details of the movements would decrease their statistical value because of the slow rate of fading.

## 3. The Nature of the Irregularities in the Ionosphere

In the process of reflection from an irregular surface the amplitude and phase of an incident wave are locally altered and diffraction theory may be used to derive the statistical properties of the field which will be observed at the ground. Under certain circumstances the nature of the irregularities may be deduced from the observed fluctuations of the field and in this part of the paper such an examination of the fluctuation data is described.

### 3.1. *Theoretical considerations*

Let the reflection process impose a complex phase modulation of r.m.s. value $\phi_m$ on the reflected wave. This can be written as $\phi_m{}^2 = \mathscr{A}_R{}^2 + \phi_R{}^2$, where $\phi_R$ is the component due to phase modulation, and $\mathscr{A}_R$ the component due to amplitude modulation at the reflector. In general, as the wavefront travels away from the ionosphere, the amount of amplitude and phase modulation, $\mathscr{A}$ and $\phi$, will change, but $\phi_m$ is invariant; i.e.

$$\mathscr{A}^2 + \phi^2 = \phi_m{}^2 = \mathscr{A}_R{}^2 + \phi_R{}^2. \tag{4}$$

In the following discussion distances in the plane of the reflection are denoted by $L$, those in the emerging wave front by $D$, and those in the wave front at the ground, where observations are made, by $d$. The spatial correlation coefficient of the irregularities of the wavefront in each plane is assumed to be Gaussian, of form $\rho(d) = \exp[-d^2/2d_0{}^2]$, and $L_0$, $D_0$, $d_0$, are the r.m.s. distances in their respective planes at which the coefficient falls to 0·61 and are used to define the structure size. The Gaussian form is chosen for ease of manipulation, not because it is necessarily observed in practice. When phase or amplitude is correlated, the fact is indicated by a suffix e.g.

$$\rho_A(d),\ \rho_\phi(d).$$

It can be shown that if $\phi_m{}^2 \ll 1$, the signal is shallowly modulated and the wave leaving the reflector has the same structure as the reflector,

i.e. $$D_0 = L_0. \tag{5}$$

If $\phi_m{}^2 \gg 1$, the signal is said to be deeply modulated and in this case

$$D_0 = L_0/\phi_m. \tag{6}$$

Bowhill (1957) showed how the structure size at the ground, $d_0$, and the magnitude of the spatial amplitude and phase fluctuations there, depend on a parameter

$$\alpha = \lambda z/\pi D_0{}^2 \tag{7}$$

when the reflector changes only the amplitude or only the phase of the incident wave (i.e. $\phi_R = 0$ or $\mathscr{A}_R = 0$).

Here, $\lambda$ = wavelength of radiation,

$z$ = distance from reflector to receiver.

If the phase only is changed at the reflector amplitude and phase variations

build up in the receding wave, as $z$, and hence $\alpha$, increases in accordance with the following expressions;

$$\mathscr{A}^2 = \frac{\alpha^2}{2(1+\alpha^2)}\phi_m{}^2; \qquad \phi^2 = \frac{2+\alpha^2}{2(1+\alpha^2)}\phi_m{}^2$$

$$\frac{\phi^2}{\mathscr{A}^2} = \frac{2+\alpha^2}{\alpha^2}. \tag{8}$$

Eventually, $\phi^2 = \mathscr{A}^2 = \frac{1}{2}\phi_m{}^2$, when $z = \infty$.
The structure size at the ground changes less markedly with $\alpha$, as shown by the values in Table 2.

Table 2. Change of phase and amplitude structure size with parameter $\alpha$ for a phase modulating reflector

| $\alpha$ | Plane wave | | Point source | |
|---|---|---|---|---|
| | $d_0(\phi)$ | $d_0(\mathscr{A})$ | $d_0(\phi)$ | $d_0(\mathscr{A})$ |
| 0 | $D_0$ | $0{\cdot}58D_0$ | $2D_0$ | $1{\cdot}42D_0$ |
| 1 | $1{\cdot}23D_0$ | $1{\cdot}71D_0$ | $2{\cdot}46D_0$ | $1{\cdot}42D_0$ |
| $\infty$ | $D_0$ | $D_0$ | $2D_0$ | $2D_0$ |

$$\alpha = \frac{\lambda z}{\pi D_0{}^2} : \frac{d_0}{D_0} = \frac{\sqrt{[(1+\alpha^2)(2+\alpha^2)]}}{\alpha^4+\alpha^2+2} \text{ for phase.}$$

$$\sqrt{\left(\frac{1+\alpha^2}{3+\alpha^2}\right)} \text{ for amplitude.}$$

When the reflector is midway between transmitter and receiver, the structure size of the irregularities at the receiver are twice as large as when the source is at infinity. Thus, if the reflector is phase-modulating, equation (8) and Table 2 show that the amplitude deviations are, in general, smaller and with finer structure, than the phase deviations. For a reflector which changes only the amplitude, the situation is reversed and the phase fluctuates least.

Bowhill pointed out that from low frequency measurements made near the reflector at nearly vertical incidence, where $\alpha$ is small, and under conditions in which $D_0$ is large, the intrinsic nature of the modulation could be determined. The data from a series of short-range measurements of phase and amplitude have been examined with this objective in view and the results are given in the following sections.

The possibility of a simple height variation of the reflecting layer being sufficient to explain the observations has also been investigated using, in addition, the medium range measurements described in Section 2.1. The variation of path length is given by $2\Delta h \cos\psi$, where $h$ is the mean ionospheric height, $\Delta h$ the r.m.s. height variation and $\psi$ the angle of incidence at the layer.

Hence $$\phi_R = \frac{4\pi\Delta h \cos\psi}{\lambda} = \phi_m \qquad (\text{since } \mathscr{A}_R = 0). \tag{9}$$



VOL. 20 1961

$$\text{If} \quad \lambda > 4\pi\Delta h \cos\psi,\ \phi_m < 1 \quad \therefore D_0 = L_0 \quad \text{and} \quad d_0 = L_0 B. \tag{10}$$

$$\text{If} \quad \lambda < 4\pi\Delta h \cos\psi,\ \phi_m > 1 \quad \therefore D_0 = \frac{L_0}{\phi_m} \quad \text{and} \quad d_0 = \frac{L_0 B\lambda}{4\pi\Delta h \cos\psi} \tag{11}$$

where $B$ is BOWHILL's factor, $d_0/D_0$ (see Table 2).

3.2. *Ionospheric characteristics from short-range measurements*

During 1956, the phase and amplitude of the signals from GBR, the 16 kc/s transmitter at Rugby, were measured on several occasions at and near Slough at a range of 108 km from the transmitter (PRESSEY *et al.*, 1960). Attention was confined to that part of the ionospheric wave which is horizontally polarized. Measurements by day and night were made with two receiving equipments working

Table 3. Ionospheric characteristics deduced from 16 kc/s observations at a range of 108 km

| | Mean $\phi$ | Mean $\mathscr{A}$ | $\phi^2/\mathscr{A}^2$ | $\alpha$ | $D_0$(km) |
|---|---|---|---|---|---|
| Day | 0·11 ± 0·02 | 0·056 ± 0·001 | 3·9 ± 1·6 | 0·84 ± 0·24 | 17·9 ± 2·6 |
| Night | 0·52 ± 0·09 | 0·18 ± 0·02 | 0·8 ± 3·2 | 0·53 ± 0·12 | 23·6 ± 2·7 |

simultaneously and separated by between 1 km and 75 km. Two independent estimates of the magnitudes of the phase and amplitude fluctuations were thus obtained simultaneously, and the spatial correlation coefficient for both amplitude and phase was derived.

The standard deviation for four night periods was

$$\phi = 0{\cdot}52 \pm 0{\cdot}09;\ \mathscr{A} = 0{\cdot}18 \pm 0{\cdot}02$$

and for two day periods was

$$\phi = 0{\cdot}11 \pm 0{\cdot}02;\ \mathscr{A} = 0{\cdot}056 \pm 0{\cdot}001.$$

The observation periods varied between 1·7 and 6 hr in duration; the night figures represent a total of 27·3 hr of observation, and the day ones a total of 9·3 hr. In each case $\mathscr{A}$ is significantly smaller than $\phi$. Moreover the fading times and correlation distances tended to be greater for phase, though the actual figures varied greatly from day to day. These facts suggest initial phase, rather than amplitude, variations.

Let us assume such a reflecting surface, changing the phase but not the amplitude of the wave. As $\phi$ and $\mathscr{A}$ are less than unity, shallow modulation conditions apply. Then we can calculate $\alpha$ from (8) and $D_0$ from (7).
Table 3 shows that the size of the irregularities is near to 20 km by both day and night, being slightly smaller by day.

From (4), we find that

$$\phi_m = 0{\cdot}12 \pm 0{\cdot}02 \text{ for day} \quad \text{and} \quad 0{\cdot}55 \pm 0{\cdot}08 \text{ for night.}$$

Assuming that the irregularities are produced by height fluctuations, then using

(9) and taking the mean height of the reflecting layer as 85 km at night and 70 km by day we obtain

$$\Delta h = 235 \pm 40 \text{ m by day} \qquad \text{and} \qquad 990 \pm 140 \text{ m by night.}$$

Applying BOWHILL's factor to the values of $D_0$ given in Table 3 we obtain the correlation distances on the ground for ranges of 108 km, and 1000 km shown in Table 4.

Table 4. Amplitude and phase correlation distances observed from a point transmitter

| | Range = 108 km, $f$ = 16·0 kc/s | | | Range = 1000 km, $f$ = 17·2 kc/s | |
|---|---|---|---|---|---|
| | $d_0(\phi)/d_0(\mathscr{A})$ | $d_0(\mathscr{A})$ (km) | $d_0(\phi)$ (km) | $d_0(\mathscr{A})$ (km) | $d_0(\phi)$ (km) |
| Day | 1·8 | 24·8 ± 5·2 | 43·0 ± 8·0 | | |
| Night | 1·8 | 29·6 ± 4·0 | 53·6 ± 7·6 | 43 ± 5·7 | 52 ± 7·4 |

The value of 43 km for $d_0(\mathscr{A})$ at a range of 1000 km may be compared with the measured value of 46 km (Section 2.3). No comparisons at the greater distances can be made, however, since BOWHILL's method does not apply where multi-hop propagation is involved.

### 3.3. *Ionospheric height variations from medium range measurements*

Variations in the height of the reflecting layer have also been deduced from the Varberg night-time measurements described in Section 2.1.

The correlation distance $d_0(\mathscr{A})$ given by the analysis of these measurements was 46 ± 8 km and the magnitude of the amplitude fluctuation of the ionospheric wave $a/A$ was 0·1 ± 0·004. It was also shown that the phase fluctuations and the corresponding amplitude fluctuations were approximately equal. Thus we can write $\mathscr{A} = \phi = 0{\cdot}1 \pm 0{\cdot}004$ and putting $\mathscr{A}_R = 0$ calculate from (4) a value of $\phi_R = 0{\cdot}14 \pm 0{\cdot}006$. Taking $h$ = 85 km and $\psi = 78\frac{1}{2}°$, (9) gives a value for $\Delta h$ of 970 ± 46 km which agrees very closely with that obtained from the short range measurements.

## 4. CONCLUSIONS

Observations at 17·2 kc/s show that the total amplitude of the signal received 1000 km from the transmitter has random fluctuations of 0·153 ± 0·007 of the mean amplitude, the correlation pattern is 46 ± 8 km wide and the fading time is 18 ± 4 min. At 15·50 kc/s, and 5800 km range, the fluctuations are 0·118 ± 0·006 of the mean, with a pattern 115 ± 12 km wide and a fading time of 26 ± 4 min. Each receiving point was far enough from the ionosphere for a random signal, whose phase and amplitude characteristics are similar to be seen. The standard deviation of phase deduced from the amplitude measurements is 8·8° at 1000 km and 6·8° at 5800 km. Variations of the observed ionospheric properties from day to day were small and within the probable variation attributable to the limited observing periods. They showed no dependence on season or other geophysical phenomena. The fading is produced by irregular changes of structure rather than by steady drifts in the ionosphere.

A model of the lower ionosphere with constant reflection coefficient but random undulations of reflection height can explain the observed facts of v.l.f. fading. The r.m.s. height variation is just less than 1 km at night, when the near height is assumed to be 85 km, and about a quarter of this in the daytime, when the mean height is 70 km. The spatial correlation coefficient of the amplitude fluctuations at the ionosphere falls to 0·61 in about 23 km at night. The same value for the correlation distance was obtained by MERCIER (1959) from his analysis of the components of the ionospheric wave received from GBR, at Rugby at nearly vertical incidence.

*Acknowledgements*—The author wishes to express his appreciation of the assistance given by a number of Universities and Government Establishments who provided accommodation for the receiving equipment and, in some cases, assisted in the maintenance of it. Special thanks are due to the Meteorological Office who provided facilities for the installation of equipment in their Ocean Weather ships and to the members of their staff who operated it.

The author also gratefully acknowledges the assistance given by his colleagues at the Radio Research Station in the analysis of the observations and the preparation of the paper.

This work forms part of the programme of the Radio Research Board of Department of Scientific and Industrial Research and is published by permission of the Director of Radio Research.

VOL.
20
1961

## REFERENCES


| | | |
|---|---|---|
| BOOKER H. G., RATCLIFFE J. A. and SHINN D. H. | 1950 | *Phil. Trans.* A **242,** 579. |
| BOWHILL S. A. | 1956 | *J. Atmosph. Terr. Phys.* **8,** 129 |
| BOWHILL S. A. | 1957 | *J. Atmosph. Terr. Phys.* **11,** 91. |
| BRAMLEY E. N. | 1951 | *Proc. Instn Elect. Engrs* III **98,** 98. |
| BRIGGS B. H., PHILLIPS G. J. and SHINN D. H. | 1950 | *Proc. Phys. Soc. Lond.* B **63,** 106. |
| HARWOOD J. | 1953 | *Proc. Instn Elect. Engrs* IV **101,** 183. |
| MERCIER R. P. | 1959 | *Phil. Mag.* **4,** 763. |
| PRESSEY B. G., ASHWELL G. E. and HARGREAVES J. K. | 1960 | *Proc. Instn Elect. Engrs* In press. |
| RATCLIFFE J. A. | 1956 | *Rep. Progr. Phys.* **19,** 188. |
| WAIT J. R. and MURPHY A. | 1956 | *Goefis. Pur. Appl.* **35,** 61. |

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 167 to 176. Pergamon Press Ltd. Printed in Northern Ireland

# On the influence of horizontal motion of the neutral air on the diffusion equation of the *F*-region

J. P. DOUGHERTY
Cavendish Laboratory, Cambridge

(*Received* 21 *September* 1960)

**Abstract**—The motion of the ionization in the *F*-region is calculated, taking into account diffusion, electrostatic fields, the magnetic field, and in addition, an arbitrary horizontal "wind" in the neutral air. Horizontal variations of all quantities are neglected. The vertical component of velocity is used to obtain a "diffusion equation" in the usual way.

The reaction of the ions on the neutral air (through collisions) is then studied. An equilibrium can be set up in a time of the order of half an hour, and if this does happen, the vertical velocity induced by the electrostatic field is drastically reduced compared with its value when the neutral air is at rest; the horizontal drift velocity, however, is increased. Further, the factor $\sin^2 I$ modifying the diffusion coefficient is removed. Thus the diffusion equation is altered considerably.

An attempt to discuss the world-wide consequences of this effect shows the need for further study of possible causes of motion of the neutral air.

## 1. INTRODUCTION

MANY writers have considered the equations governing the formation and decay of the ionospheric *F*-region in an attempt to understand the massive amount of data that has been collected by radio sounding. It is well known that the electron density $N(h)$ at a given height $h$ is not determined by the rates of production, $q$, and loss, $L$, at that height alone, because vertical transport of ions and electrons is also effective. The basic theory was first given by FERRARO (1945), who concluded that the effect was negligible; however, the air density is now known to be less than was thought in 1945, so that more recent writers (e.g. DUNGEY, 1956; YONEZAWA, 1958, and earlier references cited there) have been obliged to include a transport term in their equations.

Since the earth's magnetic field is in general inclined to the vertical, it is difficult to discuss the vertical velocity of charged particles without writing down the complete equations of motion. But the horizontal velocity is also of interest, since horizontal "drifts" of ionization are observed by routine methods (BRIGGS and SPENCER, 1954). The forces giving rise to motion of the ionization include partial pressure (or concentration) gradient, gravity, friction due to collisions with neutral particles, and electromagnetic forces due to potential fields generated at the dynamo level and transmitted to the *F*-region (MARTYN, 1953, 1955)—the description "ionospheric motor" has been applied to the *F*-region in this connexion. Finally, if the neutral air is moving, it tends to drag the ionization with it. Little attention seems to have been given to this last effect. Some authors tacitly assume that the neutral air is at rest; others include an arbitrary vertical "drift velocity", of unspecified cause, in their diffusion equation.

In the present paper we consider first the general equations of motion for both the charged and neutral species taking into account all the above effects. Then the consequences of two simple assumptions about the neutral gas are compared.

The differences are striking, showing that we need further investigation into the causes of motion of the neutral gas before detailed work on the diffusion equation can proceed unambiguously.

## 2. Basic Assumptions

Neglecting the curvature of the earth, we take axes $Ox$, $Oy$ and $Oz$ pointing south, east and upwards respectively. Thus the magnetic field has components

$$\mathbf{B} = B(-\cos I, \mathrm{O}, -\sin I) \tag{1}$$

where $I$ is the dip angle, positive in the Northern hemisphere.

We shall repeatedly make use of the assumption that the ionosphere is horizontally stratified, so that all the variables depend only upon $z$ and the time, $t$. Possible sources of error arising from the neglect of horizontal gradients will be considered later.

So far as the electromagnetic field is concerned, we shall ignore the magnetic field produced by currents in the ionosphere compared with the earth's main field. Electric fields induced by the variation of these magnetic fields are also ignored; in other words, we consider only the mean which remains after averaging over hydromagnetic oscillations. As usual, the Poisson equation merely evaluates the small difference $N_i - N_e$ between the ion and electron densities, which is ignored everywhere else in the theory. (We write $N$ for both $N_i$ and $N_e$.) So all that survives of Maxwell's equations is

$$\operatorname{curl} \mathbf{E} = 0. \tag{2}$$

But as already explained, variations with respect the $x$ and $y$ are to be ignored. Thus one component of (2) is identically satisfied, and the others reduce to $\partial E_x/\partial z = \partial E_y/\partial z = 0$, i.e. $E_x$ and $E_y$ are independent of $z$. We shall suppose, as an idealization, that $E_x$ and $E_y$ are given at some "base level" (the dynamo region) as a boundary condition. Hence only $E_z$ appears as a variable in our problem.

The dynamical equations consist of the usual continuity and momentum equations for each kind of particle. The effect of collisions may be represented by the usual frictional term. The present writer finds it more convenient in this kind of work to treat each gas separately in this way, rather than to write down generalized Ohm's law equations with conductivities $\sigma_0$, $\sigma_1$, $\sigma_2$, $\sigma_3$ (the usual notation). The two methods are of course fundamentally equivalent. Let us take the neutral and charged particles in turn.

## 3. Motions of the Neutral Air

The equation of motion is

$$\frac{d\mathbf{U}}{dt} = -\frac{1}{\rho}\nabla p + \mathbf{g} + \zeta\nabla^2\mathbf{U} + \frac{Nm_i\nu_i}{N_n m_n}(\mathbf{u}_i - \mathbf{U}) \tag{3}$$

where $\mathbf{U} = (U,V,W)$ is the hydrodynamic velocity, $p$ is the pressure, $\rho = N_n m_n$ the mass density, $\mathbf{g} = -(0,0,g)$, $\zeta$ is the kinematical viscosity, and $\nu_i$ the collision frequency of ions (velocity $\mathbf{u}_i$) with neutral molecules. (The frictional term due to electron collisions is quite negligible by comparison with that due to ions.)



VOL. 20 1961

The vertical component of (3) is primarily the equation of static equilibrium, $\partial p/\partial z = -g\rho$. Clearly, in a steady and strictly horizontally stratified atmosphere, $W$ must be zero (there being no source or sink of air at infinity), so the effect of frictional drag is only to modify the equilibrium density distribution. To see that this modification is slight, we note that at 300 km height $N/N_n$ may be about 1/300 and $\nu_i = 0{\cdot}25\ \text{sec}^{-1}$. If we take 100 m/sec as a typical ionospheric velocity to insert in $(w_i - W)$ we find that the drag term corresponds to an acceleration of about $0{\cdot}1\ \text{m sec}^{-2}$, whereas $g \approx 10\ \text{m sec}^{-2}$. Thus a 1 per cent modification of the density distribution can cancel the frictional term. We neglect such modifications: this is equivalent to regarding the neutral air as unmovable *in the vertical direction*.

For the horizontal part of (3), $\mathbf{g} = 0$ and $\nabla p = 0$, so that in a steady solution one merely has the last two terms. It is quite likely that (even disregarding the charged particles) there exist small horizontal pressure gradients due to thermal effects or perturbing gravitational fields causing tidal motions. It is not known how important these are, and whether they are likely to be very small or large compared with the drag of the ions. For the present, let us regard $U, V$ as *arbitrary* and examine the equations for the charged particles. We can then return to this question having estimated $\mathbf{u}_i$ for any $U, V$.

## 4. Motions of the Plasma

The momentum equations for ions and electrons are

$$\left.\begin{aligned} m_i\nu_i(\mathbf{u}_i - \mathbf{U}) + m_e\nu_{ei}(\mathbf{u}_i - \mathbf{u}_e) - e\mathbf{u}_i \times \mathbf{B} &= e\mathbf{E} - \frac{K}{N}\nabla(TN) + m_i\mathbf{g} \\ m_e\nu_e(\mathbf{u}_e - \mathbf{U}) + m_e\nu_{ei}(\mathbf{u}_e - \mathbf{u}_i) + e\mathbf{u}_e \times \mathbf{B} &= -e\mathbf{E} - \frac{K}{N}\nabla(TN) + m_e\mathbf{g} \end{aligned}\right\} \quad (4)$$

The second term in each equation represents a collisonal drag due to electron-ion collisons with frequency $\nu_{ei}$; $K$ is the Boltzmann constant and $T$ the temperature.

Some explanation about the role of collisions in this problem should perhaps be given. In omitting the acceleration terms in equations (4), and in writing a simple "scalar pressure" term (rather than using Boltzmann's equation) and in assuming that all three gases are at the same temperature, we are tacitly assuming that the dynamics is in one sense "collision-dominated". On the other hand, we shall then proceed to ignore certain of the collisional terms in (4). That these are consistent procedures is to be explained as follows. The neglect of the acceleration terms and the use of ordinary "continuum" fluid equations require that the time between collisions $\tau = 1/\nu$ be small compared with a typical "hydrodynamic" time $\tau_h$, viz. the time in which *macroscopic* features of the flow are changing. The procedure is essentially an expansion in powers of $\tau/\tau_h$. Now $\tau_i$ (for the ions) is at the most about 5 sec and $\tau_e$ is much less, while $\tau_h$ in our applications will generally be of the order of hours or a day. On the other hand, when motion across the magnetic field is considered, the gyro frequency, $\omega$, becomes the dominant parameter and collisons cease to be important, in the sense that $\nu \ll \omega$ especially for the electrons. We shall systematically ignore terms of order $\nu_e/\omega_e$ in comparison with terms of order 1. We retain terms of order $\nu_i/\omega_i$ at first, but omit terms of order $\nu_i^2/\omega_i^2$.

No such argument applies for components parallel to the field, and we do indeed retain the frictional term for the ions. For the electrons however, we still ignore it in this problem. This point is not solely a question of comparison of terms in the differential equation—one must also bear in mind the boundary conditions. In a plasma, such as the $F$-region, if one applies an external electric field $\mathbf{E}$ parallel to $\mathbf{B}$, it is true that a large current would flow, and the magnitude of the current would depend on $\nu_e$ and $\nu_{ei}$; in such a case, the $\nu$'s can clearly *not* be ignored. But in the present problem, only the values $E_x$ and $E_y$ are given by external conditions, while $E_z$ is an unknown variable which it is part of the problem to discover. But we can reasonably suppose as a further boundary condition that large steady currents do not flow across the boundary, i.e. there is no large vertical component of current. In this case, the component parallel to $\mathbf{B}$ of the equation for electrons becomes more in the nature of a condition on $E_z$ than an equation for the velocity; in fact, it asserts that $E_z$ is adjusted so that $\mathbf{E}.\mathbf{B} \approx 0$, with a residual effect due to gravity and pressure, which as we shall see is comparatively small.

Finally, it will emerge from the analysis that, even neglecting $\nu_{ei}$, the plasma shows a marked tendency to avoid carrying any *current* in the sense that $\mathbf{u}_i - \mathbf{u}_e$ is small compared with $\mathbf{u}_i$. This provides further justification for the neglect of the term in $\nu_{ei}$, whose effect is only to increase slightly this tendency.

We can now reduce (4) to the simpler form

$$\left.\begin{aligned} m_i\nu_i(\mathbf{u}_i - \mathbf{U}) - e\mathbf{u}_i \times \mathbf{B} &= e\mathbf{E} - \frac{K}{N}\nabla(TN) + m_i\mathbf{g} \\ e\mathbf{u}_e \times \mathbf{B} &= -e\mathbf{E} - \frac{K}{N}\nabla(TN) \end{aligned}\right\} \tag{5}$$

where we also ignore the gravitation of the electrons.

The equations of continuity in this case are

$$\frac{\partial N}{\partial t} + \frac{\partial}{\partial z}(Nw_{i,e}) = q - L. \tag{6}$$

For boundary conditions, we have firstly that $E_x$ and $E_y$ are given, hence are known at all heights, and also that $(Nw) \to 0$ as $z \to \infty$, (there is no external "source" of ionization). Finally, a condition on $N$ at the level where $E_x$ and $E_y$ are imposed, should be given; but the problem is hardly changed if we suppose this level to be at great distance, idealize it to $z \to -\infty$, and impose the condition $N \to 0$ there.

The condition at $z \to \infty$ together with (6) show that we must have $w_i = w_e$ at all levels (no vertical current), so write $w$ for both of them. We thus lose one variable, but (6) reduces to one equation. Then (5) and (6) give seven conditions on $u_{i,e}, v_{i,e}, w, N$ and $E_z$. So one first solves (5) for the velocities and $E_z$ since it is purely algebraic (i.e. not differential) in these variables.

The scalar product of the second equation (5) with $\mathbf{B}$ gives

$$E_z = -E_x \cot I - \frac{K}{Ne}\frac{\partial(TN)}{\partial z} \tag{7}$$



VOL. 20 1961

showing the slight departure from $\mathbf{E}.\mathbf{B} = 0$ due to pressure.

If (7) is used to solve the remaining five independent equations of (5) for $u_{i,e}, v_{i,e}$ and $w$ in terms of $E_x, E_y, U, V$, and $\partial N/\partial z$, it is convenient to define, merely as an abbrevation,

$$eE_p = \frac{2K}{N}\frac{\partial(TN)}{\partial z} + m_i g \tag{8}$$

and also $$\alpha = 1 + (\nu_i/\omega_i)^2, \qquad \beta = 1 + (\nu_i/\omega_i \sin I)^2 \tag{9}$$

and the result may be written

$$\left.\begin{aligned}
\alpha u_i &= \frac{\nu_i}{\omega_i}\frac{E_x}{B} - \frac{E_y}{B}\sin I + U\left(\cos^2 I + \frac{\nu_i^2}{\omega_i^2}\right) - \frac{\nu_i}{\omega_i} V \sin I \\
&\quad - \frac{e}{m_i \nu_i} E_p \sin I \cos I \\
\alpha u_e &= -\frac{\nu_i}{\omega_i}\frac{E_x}{B}\cot^2 I - \beta\frac{E_y}{B}\sin I + U\cos^2 I + \frac{\nu_i}{\omega_i} V \cos I \cot I \\
&\quad - \frac{\beta e}{m_i \nu_i} E_p \sin I \cos I \\
\alpha v_i &= \frac{E_x}{B}\operatorname{cosec} I + \frac{\nu_i}{\omega_i}\frac{E_y}{B} + \frac{\nu_i}{\omega_i} U \sin I + \frac{\nu_i^2}{\omega_i^2} V + \frac{E_p}{B}\cos I \\
v_e &= \frac{E_x}{B}\operatorname{cosec} I \\
\alpha w &= -\frac{\nu_i}{\omega_i}\frac{E_x}{B}\cot I + \frac{E_y}{B}\cos I + U \sin I \cos I + \frac{\nu_i}{\omega_i} V \cos I \\
&\quad - \frac{\beta e}{m_i \nu_i} E_p \sin^2 I
\end{aligned}\right\} \tag{10}$$

These equations result *exactly* from (5); if we now carry out our intention to neglect $(\nu_i/\omega_i)^2$ we put $\alpha = 1$, but we can only put $\beta = 1$ if

$$\sin I \gg \nu_i/\omega_i. \tag{11}$$

We must therefore, exclude a region close to the magnetic equator from our considerations. There are in fact very peculiar effects known observationally within a few degrees of the equator, and our basic assumption of horizontal stratification probably fails there also, especially at great heights.

Using (6), (8) and the expression for $w$ in (10), (with $\alpha = \beta = 1$), we reach a "diffusion equation":

$$\frac{\partial N}{\partial t} = q - L - \frac{\partial}{\partial z}[N w_d] + \sin^2 I \frac{\partial}{\partial z}\left[\frac{1}{\nu_i}\left(\frac{2K}{m_i}\frac{\partial(TN)}{\partial z} + Ng\right)\right]. \tag{12}$$

Here

$$w_d = -\frac{\nu_i}{\omega_i}\frac{E_x}{B}\cot I + \frac{E_y}{B}\cos I + U \sin I \cos I + \frac{\nu_i}{\omega_i} V \cos I \tag{13}$$

is our expression for what has been called the "vertical drift due to effects other than diffusion". Equation (12) has the usual form, since the ambipolar diffusion coefficient (on a simple theory) is $2KT/m_i\nu_i$.

Let us now look more closely at equations (10) (after putting $\alpha = \beta = 1$ and neglecting $\nu_i^2/\omega_i^2$) and (13). We notice that $E_x/B$, $E_y/B$, $U$, $V$, each appear in terms with either a simple trigonometric coefficient (of order unity), or with a coefficient proportional to $\nu_i/\omega_i$. $E_p$, on the other hand, appears with coefficients $e/m_i\nu_i$ $(= \omega_i/\nu_i B)$ and $1/B$. So we must appeal to observation to estimate the relative importance of $E_p$ and $E_x$, $E_y$, $U$, $V$. We may reasonably suppose that $E_x/B$, $E_y/B$, $U$, $V$, are of the same general order of magnitude as the observed velocities, namely $\approx 50$ m/sec, since $E_x$, $E_y$ are generated from velocities of this same order in the dynamo region—at any rate, they cannot be much *larger* than this, or velocities of higher order would be observed in some circumstances, through (11). $E_p$ is so far just a shorthand for the expression in (8)—physically, we can think of it as an electric field "equivalent" to the force of gravity and partial pressure per ion-electron pair. Now for a molecular weight of 27, $m_i g/e \approx 2{\cdot}8 \times 10^{-6}$ V/m. With $T = 1000°$ we have $KT/e = 0{\cdot}09$ V, and estimating $d(\log NT)/dz$ as not greater than $10^{-5}$ m$^{-1}$ we find $(2K/Ne)[\partial(NT)/\partial z] \approx 2 \times 10^{-6}$ V/m. Thus the two parts of $E_p$ are of the same order and $E_p \approx 5 \times 10^{-6}$ might be typical. This gives $E_p/B \approx 0{\cdot}1$ m/sec, which is 500 times smaller than a typical velocity. But $eE_p/m_i\nu_i$ would be a typical velocity at $\approx 270$ km height. Hence we see that we commit no grave error in regarding $E_p$ as of order $\nu_i/\omega_i$ times $E_x$, $E_y$, $BU$, $BV$ at most heights, though at the lower part of the region, it would in fact be rather smaller even than this.

VOL. 20 1961

What may be called the "first approximation" to (10) and (13), in which terms of order $\nu_i/\omega_i$ are omitted, may therefore be written

$$\left.\begin{aligned} u_i = u_e &= -\frac{E_y}{B}\sin I + U\cos^2 I - \frac{e}{m_i\nu_i}E_p \sin I\cos I \\ v_i = v_e &= \frac{E_x}{B}\operatorname{cosec} I \\ w &= w_d - \frac{e}{m_i\nu_i}E_p \sin^2 I \\ \text{where}\quad w_d &= \frac{E_y}{B}\cos I + U\sin I\cos I \end{aligned}\right\} \qquad (14)$$

We notice that these are just the equations one writes down if it is simply assumed that:

(1) The electric field produced by the boundary values $E_x$ and $E_y$ is uniform throughout space and given by making $\mathbf{E}.\mathbf{B} = 0$.

(2) The motion due to $\mathbf{E}$ is just $\mathbf{E} \times \mathbf{B}/B^2$ for both kinds of particles.

(3) The motion produced by $\mathbf{U}$ is the component of $\mathbf{U}$ along $\mathbf{B}$, i.e. $\mathbf{B}(\mathbf{U}.\mathbf{B})/B^2$.

We also note that to this order, $\mathbf{u}_i = \mathbf{u}_e$. As Martyn has pointed out, this is the sense in which the $F$-region can be called a "motor"—the plasma as a whole moves so as to prevent a current flowing. If one returns to (10) and calculates the

current $\mathbf{J} = Ne(\mathbf{u}_i - \mathbf{u}_e)$, one finds that the force $\mathbf{J} \times \mathbf{B}$ does indeed balance the frictional force density together with gravity, etc., as it should.

One of the difficulties with the use of (12) for $F$-region equilibrium analysis is that $w_d$ is hard to estimate, but can be a very important term. An assumption often made is that it is independent of height. In some circumstances this may be so, since $E_y$ is constant and $\mathbf{U}$ could conceivably be roughly constant, so that provided the height is such that $\nu_i \ll \omega_i$ (above $\approx$200 km), (14) shows $w_d$ would be constant. But one needs a fully developed dynamo theory to predict $E_x$ and $E_y$ and their diurnal behaviour, and as we have already indicated, $U$ and $V$ are problematical.

It may be asked whether one could use the observations of horizontal velocities to get further information to predict $w_d$ without dynamo theory (assuming the velocities given by irregularities are the same as $\mathbf{u}_i$ when $\mathbf{u}_i \approx \mathbf{u}_e$, as here). At heights where (14) are valid, we see that unfortunately $v$ gives only $E_x$, which is not required elsewhere; $u$ gives one combination of $E_y$ and $U$, (and at greater heights also involves $E_p$ as we have just seen, but that effect would be small at $\approx$200 km), while $w_d$ requires a different combination of $E_y$ and $U$. We note that $V$ only appears in the second approximation—a wind across the field is ineffective above the dynamo region.

Thus it seems essential to re-discuss the equation for $U$.

## 5. The Influence of Plasma Motions on the Neutral Gas

We can now reconsider the horizontal part of the equation of motion of the neutral gas, (3), knowing that for each choice of $U(z)$, $V(z)$, $E_x$, $E_y$, there is a definite plasma velocity $\mathbf{u}_i$, which is set up within a few seconds.*

Suppose at time $t = 0$ we have such a situation with $\mathbf{U} = 0$; then for horizontal motions in a perfectly stratified atmosphere,

$$\frac{\partial \mathbf{U}}{\partial t} = \frac{N m_i}{N_n m_n} \nu_i (\mathbf{u}_i - \mathbf{U}) + \text{viscous term}$$

where

$$\mathbf{u}_i = \mathbf{u}_i(U, V, E_x, E_y).$$

If then $\mathbf{u}_i$ remained constant, $\mathbf{U}$ would tend to $\mathbf{u}_i$, the difference diminishing exponentially in time like $\exp(-t/\tau_n)$, where $\tau_n$ is a time scale for the *neutrals* to reach an equilibrium velocity under the influence of ion-drag, and given by $\tau_n = N_n m_n / N m_i \nu_i \approx N_n / N \nu_i$, since $m_n \approx m_i$. Now $\nu_i \propto N_n T^{1/2}$ (Cowling, 1945), so $\tau_n \propto N^{-1} T^{-1/2}$, and varies comparatively slowly through the ionosphere; 20 min would be a typical value for the $F$-region. So at first we could calculate $\mathbf{u}_i$ as if $U = V = 0$, but for time scales in excess of say 1 hr, we must include ion-drag—the equilibrium thus set up would not be the solution for $U = V = 0$.

Let us for the moment neglect viscosity, and the possibility that horizontal

* The velocity also depends on $N(z)$, so this is not to say that *equilibrium* would be set up in a few seconds; if $N(z)$ were not the equilibrium distribution $\mathbf{u}_i$ would change somewhat during the setting up of equilibrium, taking $\mathbf{U}$ and $\mathbf{E}$ as fixed.

pressure gradients may be set up. Then clearly the new equilibrium that would be set up must be found by solving

$$\left.\begin{aligned} U &= u_i(U, V, E_x, E_y) \\ V &= v_i(U, V, E_x, E_y) \end{aligned}\right\} \tag{15}$$

for $U$, $V$, using equations (10). If we use the "first approximation" (14) instead, the result is simply

$$\left.\begin{aligned} U = u_{i,e} &= -\frac{E_y}{B}\operatorname{cosec} I - \frac{e}{m_i \nu_i} E_p \cot I \\ V = v_{i,e} &= \frac{E_x}{B}\operatorname{cosec} I \end{aligned}\right\} \tag{16}$$

giving

$$\begin{aligned} w &= \frac{E_y}{B}\cos I + \sin I \cos I \left(-\frac{E_y}{B}\operatorname{cosec} I - \frac{e}{m_i \nu_i} E_p \cot I\right) - \frac{e}{m_i \nu_i} E_p \sin^2 I \\ &= -\frac{e}{m_i \nu_i} E_p. \end{aligned} \tag{17}$$

Thus it has appeared that when $U$ has adjusted itself to satisfy (15), the terms in $U$ in the expression for $w$ are just such as to cancel the effect of $E_y$, and also to remove the coefficient $\sin^2 I$ in the "diffusion" term.
Equation (12) is replaced by

$$\frac{\partial N}{\partial t} = q - L + \frac{\partial}{\partial z}\left[\frac{1}{\nu_i}\left(\frac{2K}{m_i}\frac{\partial (TN)}{\partial z} + Ng\right)\right] \tag{18}$$

which is just the equation we should have if the magnetic field were vertical or absent.

If one solves (15) to the second approximation, using (11), the results have a number of additional complicating terms of order $\nu_i/\omega_i$, which would however, only be noticeable at the lowest levels, so we will not quote the details.

The comparison of the horizontal drifts (16) with those obtained by taking the simpler assumption $U = V = 0$ in (14), is also of interest. We see that $u$ is simply multiplied by $\operatorname{cosec}^2 I$, while $v$ is not altered in this approximation. Thus $u$ is always greater if the neutrals are free to move horizontally, than if they are constrained. We must next examine the effect of viscosity on the foregoing. This means the terms $\zeta \partial^2 U/\partial z^2$, $\zeta \partial^2 V/\partial z^2$ in (3). These, if appreciable, would work to set up a different equilibrium from that given by (15), in such a way that the residual drag term proportional to $\mathbf{u}_i - \mathbf{U}$ could balance the viscous force. Fortunately, there is a substantial range of heights for which neither the $\nu_i/\omega_i$ terms, which have been ignored, nor the $\mathbf{E}_p$ terms are very important, so that $U$ and $V$ as given by (16) are practically independent of height, and viscosity is unimportant. Elsewhere, its importance is rather complex to estimate since one needs to know the second derivative of $E_p/\nu_i$ with respect to height, and that requires the *solution* of the diffusion equation (18) rather than its mere formulation, which in turn demands detailed knowledge of $q(z)$, $L(N,z)$, $T(z)$ and $\nu_i(z)$.



VOL. 20 1961

## 6. World-Wide Considerations

All the above analysis leaned heavily on the assumption that all quantities were strictly independent of $x$ and $y$. One is prepared to admit that in reality there are horizontal gradients, though much smaller than the vertical ones. We might therefore, suppose that as a next approximation we may take the relations among $U$, $V$, $\mathbf{u}_{i,e}$, $\mathbf{E}$, $N$ that we have developed to be still correct *locally* in the neighbourhood on any point, with the quantities varying very slowly in longitude and latitude.

The danger of such an assumption lies in our inadequate knowledge of the terms in (3) so far ignored, namely effects due to horizontal density and temperature gradients. It is of interest to examine the implications of this assumption in the (admittedly unlikely) case of no temperature variations over the globe at each fixed height. For a wide range of heights we could then deduce (for motions with time scale in excess of 20 min) from (16) that $U = -(E_y/B)$ cosec $I$, $V = (E_x/B)$ cosec $I$, $W = 0$, since $E_p$ is important only at great heights. With $I,B$ varying over the globe in the usual way and $\mathbf{E}$ varying in a way which would emerge from dynamo theory we have to ask whether the continuity equation for neutrals (which is identically satisfied in a plane and horizontally stratified motion) can be satisfied on the surface of a sphere, without a build up of density variations in accordance with $(1/\rho)(D\rho/Dt) = -\text{div}\ \mathbf{U}$. If there are density variations then horizontal pressure gradients would be set up which would react back on the motion and slow down the neutral particles despite the ion drag. The time scale for this to become effective is given roughly by $[\text{div}\ \mathbf{U}]^{-1}$ where the divergence has to be calculated on the surface of a sphere. Using spherical polar co-ordinates with colatitude $\theta$ measured from the North pole and longitude $\phi$ measured eastwards, the $(\theta,\phi,r)$ directions are precisely the $(x,y,z)$ directions respectively, so we consider

$$U = U_\theta = \frac{E_\phi}{B_r}, \quad V = U_\phi = -\frac{E_\theta}{B_r} \tag{19}$$

since $B_r = -B \sin I$. Since $\mathbf{E}$ is the dynamo potential field we have

$$E_\theta = -\frac{1}{a}\frac{\partial S}{\partial \theta} \quad E_\phi = -\frac{1}{a \sin\theta}\frac{\partial S}{\partial \phi}$$

where $S$ is the potential and $a$ is the radius of the earth. Working with an idealized dipole field $B_r = B_0 \cos\theta$ where $B_0$ is the value at the pole, on $r = a$, we find

$$\begin{aligned}
\text{div}\ \mathbf{U} &= \frac{1}{a \sin\theta}\left[\frac{\partial}{\partial\theta}(U \sin\theta) + \frac{\partial V}{\partial\phi}\right] \\
&= \frac{1}{B_0 a^2 \sin\theta}\left[\frac{\partial}{\partial\theta}\left(\frac{-1}{\cos\theta}\frac{\partial S}{\partial\phi}\right) + \frac{\partial}{\partial\phi}\left(\frac{1}{\cos\theta}\frac{\partial S}{\partial\theta}\right)\right] \\
&= \frac{1}{B_0 a^2 \sin\theta}\left[\frac{\partial S}{\partial\phi}\left(\frac{-\sin\theta}{\cos^2\theta}\right)\right] = -\frac{1}{B_0 a^2}\frac{\partial S}{\partial\phi}\sec^2\theta = \frac{U}{a}\tan\theta
\end{aligned} \tag{20}$$

where we have made explicit use of the fact that $\mathbf{E}$ is a potential field. So the typical time scale just referred to becomes $a/(U \tan\theta)$, and if we estimate $U \approx 50$

m/sec as before and $a = 6400$ km, then $a/U \approx 1{\cdot}5$ days. Thus for motions on a time scale of about 1 day or more the back reaction of pressure gradients would become important in assessing the effect of ion-drag—assuming it is not masked by still more important effects due to unequal heating of air from place to place.

In addition, it should be pointed out that for times of the order of 1 day or more, the ordinary Coriolis force should be included.

## 7. Discussion

Our tentative considerations so far have at least shown clearly how important it is to consider the motion of neutral particles in discussing theoretically both the diffusion equation and the horizontal drift motions of the $F$-region. Two very simple possible assumptions, namely (i) $\mathbf{U} \equiv 0$ and (ii) $\mathbf{U} \equiv \mathbf{u}_i$ for horizontal components, led us to quite different conclusions in both respects. Assumption (i) implies that dynamo electric fields cause important vertical motions of the ionization and that the diffusion coefficient is modified by a factor $\sin^2 I$, while assumption (ii) removes both these effects, and also increases the horizontal drift velocities compared with those of assumptions (i). If there are heat sources which depend on position, or motions transmitted mechanically from lower regions, or tidal motions, the result would be different again. We still cannot estimate the importance of the drag of the ions since we do not know with what other effects it is to be compared, though the time-scales we have computed give some guidance.

To relieve this situation a new and fundamental investigation would be required, taking into account the spherical geometry, the earth's rotation and other factors just mentioned. Electrical coupling between magnetically conjugate points, via the exosphere, may also have to be considered.

*Acknowledgements*—This investigation was suggested by Mr. J. A. Ratcliffe and was carried out during the tenure of a D.S.I.R. Research Studentship.



## References


| | | |
|---|---|---|
| Briggs B. H. and Spenser M. | 1954 | *Rep. Progr. Phys.* **17,** 245. |
| Cowling T. G. | 1945 | *Proc. Roy. Soc.* A **183,** 453. |
| Dungey J. W. | 1956 | *J. Atmosph. Terr. Phys.* **9,** 90. |
| Ferraro V. C. A. | 1945 | *Terr. Magn. Atmos. Elect.* **50,** 215. |
| Martyn D. F. | 1953 | *Phil. Trans.* A **246,** 306. |
| Martyn D. F. | 1955 | *The Physics of the Ionosphere,* pp. 163, 254, 260. (Physical Society, London). |
| Yonezawa T. | 1958 | *J. Radio Res. Lab.* **5,** 165. |

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 177 to 182. Pergamon Press Ltd. Printed in Northern Ireland

# Ion–neutral reactions

M. HERTZBERG
Lockheed Missiles and Space Division, Palo Alto, California

(*Received* 27 *September* 1960)

**Abstract**—Since the rate coefficients for exoenergetic ion–neutral metatheses are of the order of $10^{-9}$ cm$^3$ sec$^{-1}$, they dominate in controlling the positive ion composition of the ionosphere. Under a reasonable set of assumptions, an equation is derived relating the degree of nitrogen dissociation to the ionic composition. From this relationship, it is concluded that the nitrogen of the *F*-region is substantially dissociated.

## 1. INTRODUCTION

IN RECENT YEARS, both the ionospheric physicist and the reaction kineticist have displayed an increasing interest in ion–neutral reactions. To the laboratory kineticist, they provided an opportunity for obtaining further insight into the fundamental mechanism of chemical reactions. For the ionospheric physicist, a knowledge of their rate coefficients is essential for a complete understanding of ionospheric processes. Their significance in the ionosphere is discussed in Section 2; their significance to the kineticist is discussed in Section 3.

## 2. IONOSPHERE

The following ion–neutral reactions control the composition of the positive ions of the earth's ionosphere:

$$O^+ + N_2 \rightarrow NO^+ + N + 1{\cdot}114\ \text{eV} \quad (1)$$

$$O^+ + O_2 \rightarrow O_2^+ + O + 1{\cdot}53 \quad (2)$$

$$N^+ + O_2 \rightarrow NO^+ + O + 6{\cdot}70 \quad (3)$$

$$N^+ + O_2 \rightarrow NO + O^+ + 2{\cdot}32 \quad (4)$$

$$O_2^+ + N \rightarrow NO + O^+ - 0{\cdot}143 \quad (5)$$

$$O_2^+ + N \rightarrow NO^+ + O + 4{\cdot}23 \quad (6)$$

$$O_2^+ + N_2 \rightarrow NO^+ + NO + 0{\cdot}99 \quad (7)$$

$$N_2^+ + O \rightarrow NO^+ + N + 3{\cdot}10 \quad (8)$$

$$N_2^+ + N \rightarrow N_2 + N^+ + 1{\cdot}04 \quad (9)$$

$$N_2^+ + O_2 \rightarrow NO^+ + NO + 4{\cdot}50 \quad (10)$$

$$O_2^+ + N(^2D) \rightarrow NO + O^+ + 2{\cdot}23 \quad (11)$$

$$NO^+ + N(^2D) \rightarrow N_2 + O^+ + 1{\cdot}24 \quad (12)$$

These reactions "shuffle" or "scramble" the ionic species more rapidly than they can be created by photoionization, or destroyed by recombination; hence they dominate in controlling the ionic composition. Not only do they convert atomic to molecular ions, but they can reverse this process in the presence of atomic species (KRASSOVSKY, 1957; HERTZBERG, 1959a, b).

Neglecting ionization and recombination, the rate of change of concentration of atomic oxygen ions is given by the equation:

$$-\frac{d(O^+)}{dt} = -k_1(O^+)(N_2) - k_2(O^+)(O_2) + k_4(N^+)(O_2) + k_5(O_2^+)(N) + k_{11}(O_2^+)(N^2D) + k_{12}(NO^+)(N^2D) . \quad (13)$$

For steady-state conditions, the following assumptions are made:

(a) the rate coefficients for the exoenergetic ion-neutral reactions are equal,
(b) the rate coefficient for (5) is characterized by an activation energy just equal to its endoenergeticity,
(c) reaction (4) may be neglected since the ratio $(N^+)/(O^+)$ is only a few per cent and $(O_2) \ll (N_2)$.

With these assumptions, the following equation results:

$$\frac{(N)}{(N_2) + (O_2)} = \frac{(O^+)}{(O_2^+)} \left\{ \exp\left(\frac{-0{\cdot}143}{kT}\right) + x\left[1 + \frac{(NO^+)}{(O_2^+)}\right] \right\}^{-1}. \quad (14)$$

The factor $x$ is the ratio of $N(^2D)$ atoms to ground state $N(^4S)$ atoms. Rocket measurements of the $(O^+)/(O_2^+)$ ratio (JOHNSON *et al.*, 1958; ISTOMIN, 1959) are available. These have been used to calculate the degrees of dissociation given in Fig. 1. Curve A applies to the night flight NN 3.17 assuming $x = 0$. Curve B applies to the day flight NN 3.19, assuming $x = 0{\cdot}1$. In view of the gross assumptions made, the agreement is significant only insofar as it indicates a larger value of $x$ for daytime than for night-time conditions. A reaction entirely analogous to (1) is a logical choice for the source of $N(^2D)$ atoms, namely:

$$O^+(^2D) + N_2 \rightarrow NO^+ + N(^2D). \quad (15)$$

Since existing evidence (see Section 3) indicates that the rate coefficient for reaction (1) is greater than $10^{-9}$ cm$^3$ sec$^{-1}$, HERTZBERG (1958a, 1959b) reasoned that the lifetime for $(O^+)$ ions at 200 km would be too short to be consistent with the observed $(O^+)$ concentrations. As a result, it was necessary to assume that atomic oxygen ions were generated via reaction (5), and that the nitrogen of the atmosphere was substantially dissociated (HERTZBERG, 1958b). In spite of the presence of reactions (11) and (12), and even under extreme auroral conditions, the night-time concentration of $N(^2D)$ is too low to materially alter this conclusion. In analysing the same rocket measurements, YONEZAWA *et al.* (1959) concluded that the coefficients for (1) and (2) should be $5{\cdot}2 \times 10^{-10}$ and $2{\cdot}9 \times 10^{-11}$ cm$^3$ sec$^{-1}$. More recently BATES and NICOLET (1960) have taken a similar position, but their analysis required the rate constants for (1) and (2) to be of the order of $10^{-13}$ cm$^3$ sec$^{-1}$. As a result, they conclude that these reactions are characterized by an activation energy or by steric hindrance.

While reliable measurements of the rate coefficients for reactions (1) to (12) do not exist, a large variety of ion–neutral reaction rate coefficients have been measured. There is no evidence whatever for either an activation energy or for steric hindrance for an exoenergetic ion–neutral metathesis. The presence of significant concentrations of atomic nitrogen adds a negative term to the decay



VOL. 20 1961

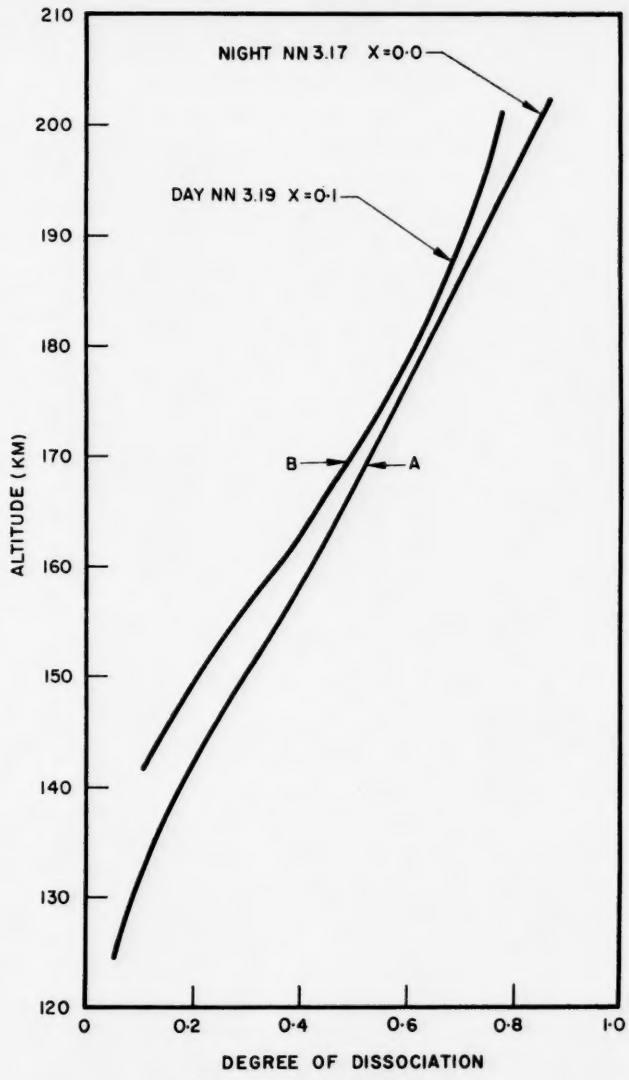


Fig. 1. Nitrogen dissociation curves.

constant for $O^+$ ions (referred to as $\gamma$ by BATES and NICOLET (1960)); the presence of this term allows the decay constant to be small even though the individual rate coefficients are as large as $10^{-9}$ cm$^3$ sec$^{-1}$.

If *neutral* nitric oxide is a significant component of the ionosphere, then the additional "scrambling" reactions that must be included are:

$$O^+ + NO \rightarrow NO^+ + O + 4{\cdot}38 \qquad (16)$$

$$N^+ + NO \rightarrow NO^+ + N + 5{\cdot}31 \qquad (17)$$

$$N^+ + NO \rightarrow N_2^+ + O + 2{\cdot}22. \qquad (18)$$

The mole fraction of nitric oxide in the 60–87 km region is less than $10^{-7}$ (JURSA *et al.*, 1959). One could justifiably neglect these reactions if this measurement could be applied to higher altitudes. This extrapolation is not possible until all sources and sinks of NO are adequately known.

In view of the rapid scrambling of positive ions via reactions (1) to (15), it is clear that *the abundance of ionic species does not bear a simple relationship to the abundance of neutrals;* that the final ionic composition will be determined by these reactions regardless of which ions are formed initially; and furthermore that reference to any reaction as resulting in a permanent loss of any constituent is clearly erroneous.

## 3. THE LABORATORY

The ion–neutral reactions (1) to (12) are intriguing examples of a variety of chemical reactions between ions and neutrals of interest both in the laboratory and the atmosphere. Reactions (1) to (3) involve the transfer of an atom from a diatomic molecule to an ion; they have been referred to as ion–atom interchange by BATES (1955a). Reaction (4) is the transfer of a negative ion fragment from a diatomic molecule to an atomic ion. Molecular ions are converted to atomic ions via reactions (5), (9), (11) and (12), which involve the transfer of a neutral atom fragment from a diatomic ion to an atom. Reactions (6) and (8) are the transfer of a positive ion fragment from a diatomic ion to a neutral atom. They are analogous to the simple proton transfer reaction

$$H_2^+ + He \rightarrow HeH^+ + H - 0{\cdot}8\ eV \qquad (19)$$

whose cross-section has recently been measured (HERTZBERG *et al.*, 1961; ORTENBURGER *et al.*, 1960). Reactions (7) and (10) involve the simultaneous destruction and formation of two bonds.

The history of ion–neutral reaction studies may be traced to the very early experiments of THOMPSON (1912, 1921). He first observed the anomalous "3" line in the spectrum of hydrogen and certain minerals subjected to electron impact bombardment. Subsequently, DEMPSTER (1916) measured the intensity of the $H_3^+$ ion as a function of hydrogen pressure and showed that it could be eliminated at low pressures. The reaction subsequently proposed (HOGNESS and LUNN, 1925; BRASEFIELD, 1928) to account for its appearance was the ion–molecule reaction:

$$H_2^+ + H_2 \rightarrow H_3^+ + H\,. \qquad (20)$$

Recently, several studies of a large variety of reaction types have been undertaken (STEVENSON and SCHISSLER, 1958; FIELD *et al.*, 1957; POTTER, 1955; GUTBIER, 1957). While it is true that they involve mainly isotopes of hydrogen, a few have involved the transfer of large groups such as the $CH_2$ radical. In all cases, the measured rate constants have been between $2 \times 10^{-10}$ and $1 \times 10^{-8}$ cm$^3$ sec$^{-1}$. There was no evidence for an activation energy for exoenergetic reactions even when the kinetic energy of the reacting ions was as low as 0·1 eV.

An essentially classical theory has been developed (EYRING *et al.*, 1936; GIOUMIOUSIS and STEVENSON, 1958; FIELD *et al.*, 1957) which agrees with experiment in order of magnitude. The rate constant is given by

$$k = 2\pi e(\alpha/\mu)^{1/2} \qquad (21)$$



VOL. 20 1961

where $e$ is the electronic charge, $\alpha$ is the polarizability of the neutral reactant, and $\mu$ is the reduced mass of the system. Typical activation energies for exoenergetic chemical metathesis rarely exceed 10 kcal/mole. An ion falling into the induced dipole field between it and a neutral reactant will certainly gain enough kinetic energy to overcome this barrier. For endoenergetic reactions an activation energy just equal to the reaction energy has been suggested (HERTZBERG, 1959b).

An accurate comparison between this classical theory and experiment is lacking, due to experimental difficulties in obtaining accurate values for the rate coefficients. These difficulties result from the presence of alternate reaction paths, each path leading to the same product ion (HERTZBERG *et al.*, 1961). In the hydrogen–helium system, two alternate paths are available for the formation of the $HeH^+$ ion. These are:

$$He^+ + H_2 \rightarrow HeH^+ + H \tag{22}$$

and

$$H_2^+ + He \rightarrow HeH^+ + H. \tag{23}$$

If a cross-section is obtained by assuming that only (22) contributes to the yield of secondary ions, then it is only an apparent cross-section which is equal to a weighted sum of the true cross-sections for both paths. By measuring the appearance potential of the $HeH^+$ ion, HERTZBERG *et al.* (1961) have proved that reaction (23) contributes substantially to the yield of $HeH^+$ ions.

For the atmospheric ion, $NO^+$, many more reaction paths are available. POTTER's value of $10^{-8}$ cm$^3$ sec$^{-1}$ for (1) was obtained from the $(NO^+)/(O^+)$ ratio assuming that all the $O^+$ ions formed moved through the source at thermal velocities. Neglecting dissociation in the source, reactions (1), (3), (7) and (10) would all contribute to the yield of $NO^+$ ions. If, as has been claimed, reactions (7) and (10) can be neglected due to an activation energy (BATES, 1955b), then only (1) and (3) contribute. As a result, POTTER's value for the cross-section would be too high by a factor of 2. On the other hand, the atomic ions formed by dissociative ionization must have had greater than thermal velocities in the source. Their residence times were therefore much smaller than thermal residence times, and in converting the cross-section to a rate constant, too low a result would be obtained. It is clear that improvement of the experimental techniques for studying the ion–neutral reactions of interest in the ionosphere is necessary.

*Acknowledgement*—The author acknowledges the very stimulating discussions he has had with Drs. F. S. JOHNSON, W. B. HANSON, D. RAPP and Mrs. I. B. ORTENBURGER.

## REFERENCES


| | | |
|---|---|---|
| BATES D. R. | 1955a | *Proc. Phys. Soc. Lond.* A **68,** 344. |
| BATES D. R. | 1955b | *Solar Eclipses and the Ionosphere,* Sect. V, p. 191. Pergamon Press, London. |
| BATES D. R. and NICOLET M. | 1960 | *J. Atmosph. Terr. Phys.* **18,** 65. |
| BRASEFIELD C. J. | 1928 | *Phys. Rev.* **31,** 52. |
| DEMPSTER A. J. | 1916 | *Phil. Mag.* **31,** 438. |
| EYRING H., HIRSCHFELDER J. O. and TAYLOR H. S. | 1936 | *J. Chem. Phys.* **4,** 479. |

| | | |
|---|---|---|
| FIELD F. H., FRANKLIN J. L. and LAMPE F. W. | 1957 | *J. Amer. Chem. Soc.* **79,** 2419. |
| GIOUMOUSIS G. and STEVENSON D. P. | 1958 | *J. Chem. Phys.* **29,** 294. |
| GUTBIER H. | 1957 | *Z. Naturf.* **129,** 499. |
| HERTZBERG M. | 1958a | Private Communication to T. YONEZAWA. |
| HERTZBERG M. | 1958b | *J. Geophys. Res.* **63,** 856. |
| HERTZBERG M. | 1959a | *J. Geophys. Res.* **64,** 1106. |
| HERTZBERG M. | 1959b | Dissertation, Stanford University. |
| HERTZBERG M., RAPP D., ORTENBURGER I. B. and BRIGLIA D. | 1961 | *J. Chem. Phys.* In press. |
| HOGNESS T. R. and LUNN E. G. | 1925 | *Phys. Rev.* **26,** 44. |
| ISTOMIN V. G. | 1959 | *Dokl. Akad. Nauk SSSR* **129,** 81. |
| JOHNSON C. Y., MEADOWS E. B. and HOLMES J. C. | 1958 | *J. Geophys. Res.* **63,** 443. |
| JURSA A. S., TAKAKA Y. and LEBLANC F. | 1959 | *Planet. Space Sci.* **1,** 161. |
| KRASSOVSKY V. I. | 1957 | *Izo. Akad. Nauk SSSR* (Ser. Geofiz) 4 504. |
| ORTENBURGER I. B., HERTZBERG M. and OGG R. A. | 1960 | *J. Chem. Phys.* **33,** 579. |
| POTTER R. F. | 1955 | *J. Chem. Phys.* **23,** 2462. |
| STEVENSON D. P. and SCHISSLER D. O. | 1958 | *J. Chem. Phys.* **29,** 282. |
| THOMSON J. J. | 1912 | *Phil. Mag.* **24,** 241. |
| THOMSON J. J. | 1921 | *Rays of Positive Electricity and their Applications to Chemical Analysis.* Longmans Green, London. |
| YONEZAWA T., TAKAHASHI H. and ARIMA Y. | 1959 | *J. Radio Res. Lab.* **6,** 21. |


VOL.
20
1961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 183 to 188. Pergamon Press Ltd. Printed in Northern Ireland

# Theoretical world curves of maximum *F2* ionization—II

J. E. C. GLIDDON
Queen Mary College, University of London, Mile End Road, London E.1
and
P. C. KENDALL
Bedford College, University of London, Regent's Park, London, N.W.1

(*Received* 6 *October* 1960)

**Abstract**—Values of $N_mF2$ have been computed from an analytical solution of the equation governing diffusion of ions in the *F2*-region. The mathematical model takes account of electron production, attachment-like recombination and a rate of diffusion which depends upon the ambient air density and also on the geomagnetic latitude. Results are plotted on a grid of geomagnetic latitude and local time and are compared with curves drawn from world-wide ionosonde records.

## INTRODUCTION

THE PRESENT PAPER is a sequel to one with a similar title by GLIDDON and KENDALL (1960a) hereafter referred to as (I). In the latter, values of the maximum electron density in the *F2*-region were calculated from an analytical solution of the equation governing vertical diffusion of ionization,

$$\partial N/\partial t = q + (-K + D\mathscr{L})N. \tag{1}$$

Here, the electron density $N$ is a function of the time $t$ and of the height $z$, $q$ is the Chapman production function, $K$ is the attachment-type coefficient of recombination, $D$ is the height-dependent coefficient of diffusion and $\mathscr{L}$ is the diffusion differential operator.

The mathematical model used in the present paper differs in two respects from that used in (I). In the first place $K$ (previously assumed for mathematical simplicity to be a constant) is now assumed to be $10^{-4} \exp[(H_2 - z)/H_1]$ sec$^{-1}$. Secondly we replace the diffusion coefficient by $D \sin^2 I$ where $I$ is the local angle of magnetic dip. The new model therefore bears a closer comparison with currently held theories of *F2*-layer formation. Details of the analytical solution have been given by GLIDDON (1959) and application to the calculation of world curves of maximum electron density follows the method outlined in (I). The parameter $\beta$ referred to in the diagrams is related to $K$ by the equation $\beta = 1{\cdot}37 \exp[(H_2 - z_0)/H_1]$, where $z_0$ is the height of maximum ion production for the overhead sun (about 180 km), $H_1$ and $H_2$ are suitable constants, of which $H_1$ is the scale height for neutral molecules. The parameter $\gamma$ is defined as $(H_1 \sin I)^2/(kD_0)$, where $D_0$ is the value of $D$ at $z_0$, $k$ the number of seconds in one radian.

## OBJECT OF PRESENT INVESTIGATION

It is well known that values of maximum electron density calculated from the equation

$$\partial N/\partial t = q - \alpha N^2 \tag{2}$$

where $\alpha$ is the (constant) coefficient of recombination, are in good agreement with observed values of critical frequency obtained for the *E* and *F1*-regions. In particular, theoretical world curves of maximum electron density compare favourably with contours representing observed values of $f_0E$ on a grid of geographic latitude and local time. It can be inferred therefore that these regions are effectively controlled by the sun's radiation. On the other hand it was pointed out by Appleton (1946) that certain features of the world-wide behaviour of the *F2*-region show strong evidence of geomagnetic control. Thus world curves of $f_0F2$ at the equinoxes show symmetry when plotted on a grid of *geomagnetic* latitude and local time but not when plotted on a grid of geographic latitude and local time. We thought it might be of interest to plot maximum electron densities computed from the mathematical model described above on a geomagnetic latitude–local time grid and to compare them with the $f_0F2$ world curves.

## Solar and Geomagnetic Control

The electron density calculated from equation (1) is a function of several parameters. For present purposes we need only explain how and to what extent we allow for the two major controlling factors, solar radiation and the geomagnetic field. As usual, it is assumed that the ionizing radiation is monochromatic and we deal only with a quiet *F2*-region. Thus ion production, besides varying with height and local time, will be also a function of geographic latitude and of the sun's declination and will therefore vary according to the length of day. The attachment coefficient also may be expected to vary with latitude and sun's declination but in the absence of observational evidence we assume merely that it decreases exponentially with altitude in the manner proposed by Ratcliffe *et al.* (1956). The geomagnetic field affects the maximum electron density through the effective diffusion coefficient, the latter decreasing to zero at the magnetic equator on account of the factor $\sin^2 I$. The distribution of ionization is also affected by electromagnetic forces which are themselves directly related to the geomagnetic field. This effect is not allowed for in our model. Thus we can claim to show the effect of the geomagnetic field only in its influence on the rate of vertical diffusion.

In considering solar and geomagnetic effects we should mention that there is a certain ambiguity or lack of definition in the representation of world values of $N_mF2$ on a grid of geomagnetic (or geographic) latitude and local time. Thus two stations on a given parallel of geomagnetic latitude may have geographic latitudes differing by as much as 23°, so they may even experience different seasons. In the accompanying diagrams we consider case (a), stations occupying the most northerly points, and case (b), stations occupying the most southerly points of parallels of geomagnetic latitude.

## Similarities and Contrasts between Theoretical and Observed World Contours

The equinoctial curves shown in Fig. 1 are not quite symmetrical about the equator for the reason stated above. As we should expect, this slight asymmetry does not appear in the corresponding curves, Fig. 2, derived by Martyn (1959) from the data of a number of observatories scattered around the world.

At about 60° N and 60° S in Fig. 2 the ionization decreases until midnight (or later) and almost at once begins to increase again. This feature is not reproduced in Fig. 1 where the ionization remains fairly constant from midnight until sunrise. A closer comparison in this respect exists between Figs. 3, 4 and MARTYN's curves for the northern solstice, Fig. 5. This would suggest that the closed contour which

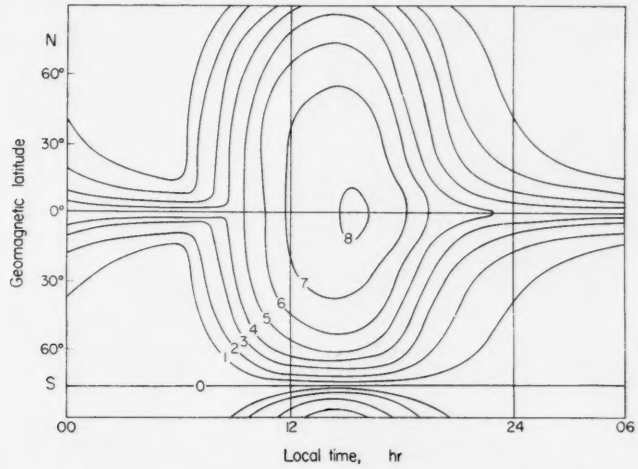


Fig. 1. Theoretical curves representing $N_m F2$ at the equinox. The effect of the geomagnetic field is allowed for in the diffusion coefficient. Case (a), $\beta = 10$, $\gamma = 20$ at lat. 45°. (Case (b) may be obtained by reflecting the curves in the equator.)

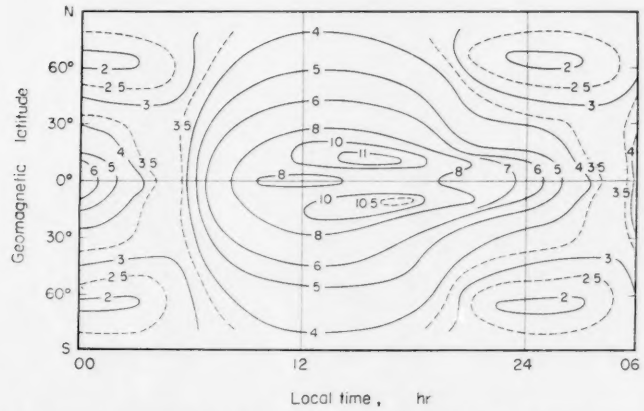


Fig. 2. World curves of $f_0 F2$ for the equinox, sunspot minimum year 1943–1944 (after MARTYN).

appears before dawn in Fig. 2 at 60° N and 60° S is a solar effect. It is possible that a similar effect would have appeared in Fig. 1 if we had been able to take the length of day to be that appropriate to the *F2*-level rather than ground level.

The equinoctial contours of Fig. 2 show very clearly that high values of $N_m$ are maintained later and later into the night as the magnetic equator is approached.



VOL. 20 1961

This effect of an "equatorial wedge" is also reproduced in the theoretical contours of Fig. 1 and is here undoubtedly due to the rate of vetical diffusion being progressively reduced towards the magnetic equator. A less prominent equatorial wedge

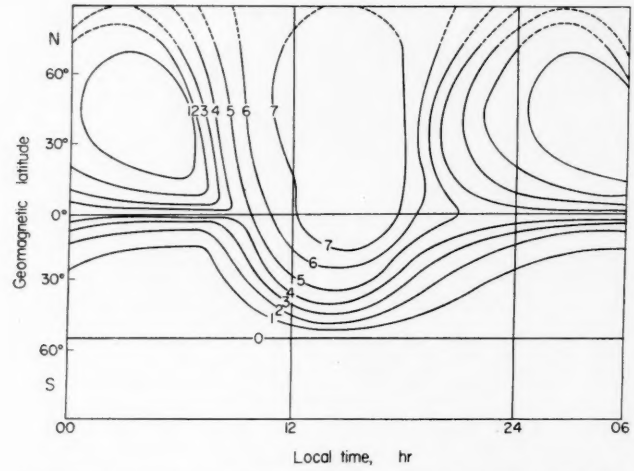


Fig. 3. Theoretical curves representing $N_mF2$ at the northern solstice when allowance is made for the effect of the geomagnetic field on vertical diffusion. Case (a), $\beta = 10, \gamma = 20$ at lat. 45°.

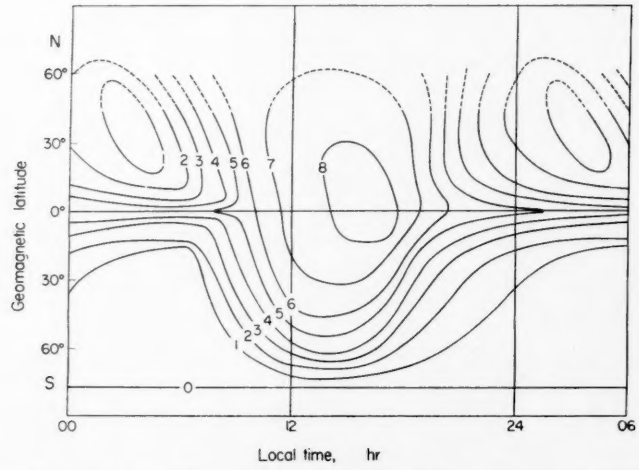


Fig. 4. Theoretical curves representing $N_mF2$ at the northern solstice when allowance is made for the effect of the geomagnetic field on vertical diffusion. Case (b), $\beta = 10, \gamma = 20$ at lat. 45°.

can be seen in Figs. 3 and 4 and may be compared with the corresponding feature in MARTYN's contours for the northern solstice. The apex of the wedge in the latter diagram (Fig. 5) will be seen to the north of the geomagnetic equator after midnight. This wedge is presumably the decaying remains of the larger anomalous peak to the north of the geomagnetic equator. In the theoretical curves, Figs. 3

and 4, the axis of the wedge after midnight is indistinguishable from the magnetic equator.

It may be mentioned here that when diffusion is ignored altogether (as at the magnetic equator) the maximum electron density occurs at infinite height. The analytical solution and computing programme worked satisfactorily, without using an excessive amount of time, for places within 1° of the magnetic equator. It will be noted from Figs. 1, 3 and 4 that near the magnetic equator high values of $N_m$ are maintained *throughout* the night and even until sunrise. It is of interest to

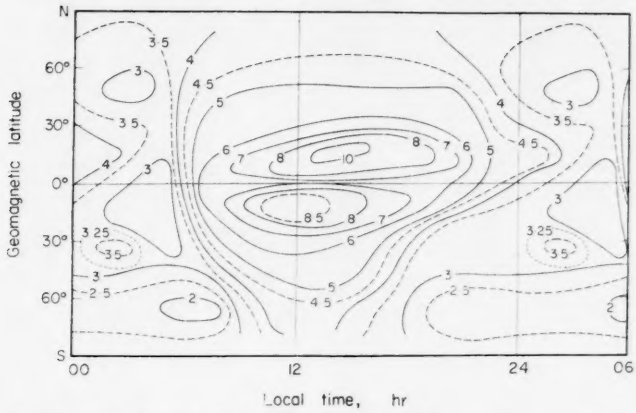


Fig. 5. World curves of $f_0F2$ for the northern solstice, sunspot minimum year 1943–1944 (after MARTYN).

note that this feature (unclosed contours surrounding the main system of closed contours) is present in MARTYN's contours for the *sunspot maximum* year 1947.

It will be noted from Fig. 1 that the theoretical contours during day-time are elongated in the north–south direction whereas the experimental curves in Fig. 2 are elongated in the direction of the time-scale. The latitude and time-scales are approximately in proportion in the two diagrams. However, Fig. 1 refers to electron density and Fig. 2 to critical frequency, and this might account for part of, but by no means the whole of, the discrepancy in configuration. It appears that our present model gives a rate of increase of $N_m$ in the morning and a rate of decrease in the evening which are too large in middle latitudes compared to the corresponding rates in the equatorial zone. We have shown in a recent paper (GLIDDON and KENDALL, 1960b) that the effect of diffusion on $N_m$ in the morning period is very slight and therefore the effect we are considering here is not likely to be explained in terms of vertical diffusion alone. It appears from our numerical results that a closer comparison could be effected if the value of the attachment-type coefficient increased north and south from the equator. This has the effect of making a given value of $N_m$ at a given latitude occur later in the morning and earlier in the evening. Such a variation could be produced, for example, if the mean scale height of the *F2*-region decreased north and south of the equator.

A striking feature of the experimental world curves is the presence of peaks of ionization to be found just after midday on either side of the magnetic equator—the



VOL. 20 1961

Appleton anomaly. These peaks do not appear in our diagrams and it seems unlikely that their formation can be explained by our present model, which takes account only of *vertical* diffusion. The importance of diffusion in the equatorial zone is in transporting ions along the nearly horizontal lines of force north and south from the equator. Mitra proposed that this effect could be large enough to cause the midday equatorial trough and the afternoon peaks of ionization. We are at present trying to construct a mathematical model to see if this proposal can be supported quantitatively.

## References


| | | |
|---|---|---|
| Appleton E. V. | 1946 | *Nature, Lond.* **157,** 691. |
| Gliddon J. E. C. | 1959 | *Quart. J. Mech.* **12,** 347. |
| Gliddon J. E. C. and Kendall P. C. | 1960a | *J. Atmosph. Terr. Phys.* **18,** 48. |
| Gliddon J. E. C. and Kendall P. C. | 1960b | *J. Geophys. Res.* **65,** 2279. |
| Martyn D. F. | 1959 | *Proc. Inst. Radio Engrs, N.Y.* **47,** 147. |
| Ratcliffe J. A., Schmerling E. R., Setty C. S. and Thomas J. L. | 1956 | *Phil. Trans.* A **248**, 621 |

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 189 to 194. Pergamon Press Ltd. Printed in Northern Ireland

# On the evaluation of the group refractive index in case of no collisions*

H. UNZ
Electrical Engineering Department, University of Kansas, Lawrence, Kansas

(*Received* 10 *October* 1960)

**Abstract**—An alternative method is shown for the evaluation of the group refractive index $\mu'$ in case of no collisions. An asymptotic form near the reflection point $X = 1$ is derived. Both of the expressions are simpler and more suitable for calculations than the ones given by SHINN and WHALE.

## 1. THE GROUP REFRACTIVE INDEX

THE group refractive index $\mu'$ is usually written in the form (RATCLIFFE, 1959, pp. 103–107)

$$\mu' = \mu + \omega \frac{d\mu}{d\omega} \tag{1}$$

where $\mu$ is the refractive index which is given for the case of no collisions by RATCLIFFE (1959, p. 53; 7):

$$\mu^2 = 1 - X\left\{1 - \frac{{Y_T}^2/2}{1-X} \pm \left[\left(\frac{{Y_T}^2/2}{1-X}\right)^2 + {Y_L}^2\right]^{1/2}\right\}^{-1} \tag{2}$$

where: $$X = \frac{{\omega_N}^2}{\omega^2}, \quad Y_L = \frac{\omega_L}{\omega}, \quad Y_T = \frac{\omega_T}{\omega}.$$

We now rewrite (1) in a form more suitable for the evaluation of $\mu'$ from $\mu$ as given in (2). By inspection, we can see that (1) may be rewritten as follows:

$$\mu' = \mu \frac{d\omega}{d\omega} + \omega \frac{d\mu}{d\omega} = \frac{d}{d\omega}(\omega\mu)$$

and we obtain the simple relationship:

$$\mu' = \frac{d}{d\omega}(\omega\mu). \tag{3}$$

We may obtain (3) directly from first principles in case of a wave motion:

$$E = E_0\, e^{i(\omega t - kx)}.$$

We have by definition (STRATTON, 1941) the phase velocity $v_p$:

$$v_p = \frac{\omega}{k}, \quad \mu = \frac{c}{v_p}, \quad \mu = c\frac{k}{\omega} \tag{4}$$

and the group velocity $v_g$:

$$v_g = \frac{d\omega}{dk}, \quad \mu' = \frac{c}{v_g}, \quad \mu' = c\frac{dk}{d\omega}. \tag{5}$$

* This work has been supported in part by Contract No. DA-44-009 Eng.-3769 at the Midwest Research Institute, Kansas City, Missouri.



VOL. 20 1961

Substituting (4) and (5) in (3) we get an identity from the definitions. Another interesting result is that if we substitute $\mu = c/v_p$, $\mu' = c/v_g$ in (3) we get:

$$\frac{1}{v_g} = \frac{d}{d\omega}\left(\frac{\omega}{v_p}\right) = \left(v_p - \omega\frac{dv_p}{d\omega}\right)(v_p^2)^{-1} = \left(1 - \frac{\omega}{v_p}\frac{dv_p}{d\omega}\right)(v_p)^{-1} \tag{6}$$

and from (6) we get:

$$v_g = v_p\left(1 - \frac{\omega}{v_p}\frac{dv_p}{d\omega}\right)^{-1}. \tag{7}$$

(7) is identical with a relationship given by Ramo and Whinnery (1953).

Relationship (3) is still not very suitable for the calculation from (2) because we have there $\mu^2$. Therefore let us write from (3):

$$2\omega\mu\mu' = 2\omega\mu\frac{d}{d\omega}(\omega\mu) = \frac{d}{d\omega}(\omega^2\mu^2) \tag{8}$$

and we get:

$$\mu' = \frac{1}{2\omega\mu}\frac{d}{d\omega}(\omega^2\mu^2) \tag{9a}$$

or

$$\mu\mu' = \frac{1}{2\omega}\frac{d}{d\omega}(\omega^2\mu^2). \tag{9b}$$

In order to show how suitable (9) is let us first take the simple case (2) with no magnetic field:

$$\mu^2 = 1 - X = 1 - \frac{{\omega_N}^2}{\omega^2} \tag{10a}$$

$$\omega^2\mu^2 = \omega^2 - {\omega_N}^2 \tag{10b}$$

$$\mu\mu' = \frac{1}{2\omega}\frac{d}{d\omega}(\omega^2\mu^2) = \frac{1}{2\omega}(2\omega) = 1 \tag{10c}$$

and we get the well-known result (Ratcliffe, 1959, pp. 103–107) by a very simple method.

## 2. The General Case

In the general case we get from (2) with the proper substitution

$$\omega^2\mu^2 = \omega^2 - {\omega_N}^2\left\{1 - \frac{{\omega_T}^2/2}{\omega^2 - {\omega_N}^2} \pm \left[\left(\frac{{\omega_T}^2/2}{\omega^2 - {\omega_N}^2}\right)^2 + \frac{{\omega_L}^2}{\omega^2}\right]^{1/2}\right\}^{-1} \tag{11}$$

and from (11) we get:

$$\begin{aligned}\frac{d}{d\omega}(\omega^2\mu^2) = 2\omega + {\omega_N}^2\Bigg\{&\frac{{\omega_T}^2\omega}{(\omega^2 - {\omega_N}^2)^2} \pm \left[-\frac{({\omega_T}^2)^2\omega}{2(\omega^2 - {\omega_N}^2)^3} - \frac{{\omega_L}^2}{\omega^3}\right] \\ &\times \left[\left(\frac{{\omega_T}^2/2}{\omega^2 - {\omega_N}^2}\right)^2 + \frac{{\omega_L}^2}{\omega^2}\right]^{-1/2}\Bigg\} \\ &\times \left\{1 - \frac{{\omega_T}^2/2}{\omega^2 - {\omega_N}^2} \pm \left[\left(\frac{{\omega_T}^2/2}{\omega^2 - {\omega_N}^2}\right)^2 + \frac{{\omega_L}^2}{\omega^2}\right]^{1/2}\right\}^{-2}\end{aligned} \tag{12}$$



VOL. 20 1961

and (12) may be rewritten as follows by using (11):

$$\frac{d}{d\omega}(\omega^2\mu^2) = 2\omega + {\omega_N}^2\left[\frac{\omega^2(1-\mu^2)}{{\omega_N}^2}\right]^2\left\{\frac{{\omega_T}^2\omega}{(\omega^2-{\omega_N}^2)^2} - \left[\frac{{\omega_T}^4\omega}{2(\omega^2-{\omega_N}^2)^3} - \frac{{\omega_L}^2}{\omega^3}\right]\right.$$
$$\left.\times\left[\frac{{\omega_N}^2}{\omega^2(1-\mu^2)} + \frac{{\omega_T}^2/2}{\omega^2-{\omega_N}^2} - 1\right]^{-1}\right\} \qquad (13)$$

and from (13) we get:

$$\frac{d}{d\omega}(\omega^2\mu^2) = 2\omega + {\omega_N}^2\left[\frac{\omega^2(1-\mu^2)}{{\omega_N}^2}\right]^2 T \qquad (14a)$$

$$T = \left\{\frac{{\omega_T}^2\omega}{(\omega^2-{\omega_N}^2)^2}\left[\frac{{\omega_N}^2}{\omega^2(1-\mu^2)} + \frac{{\omega_T}^2/2}{\omega^2-{\omega_N}^2} - 1\right] - \frac{{\omega_T}^4\omega}{2(\omega^2-{\omega_N}^2)^3} + \frac{{\omega_L}^2}{\omega^3}\right\}$$
$$\times\left\{\frac{{\omega_N}^2}{\omega^2(1-\mu^2)} + \frac{{\omega_T}^2/2}{\omega^2-{\omega_N}^2} - 1\right\}^{-1}$$
$$= \left\{\frac{{\omega_T}^2\omega}{(\omega^2-{\omega_N}^2)^2}\left[\frac{{\omega_N}^2}{\omega^2(1-\mu^2)} - 1\right] + \frac{{\omega_L}^2}{\omega^3}\right\}\left\{\frac{{\omega_N}^2}{\omega^2(1-\mu^2)} + \frac{{\omega_T}^2/2}{\omega^2-{\omega_N}^2} - 1\right\}^{-1} \qquad (14b)$$

and from (9b) we get:

$$\mu\mu' = 1 + \frac{(1-\mu^2)^2}{2}P \qquad (15a)$$

$$P = \left\{\frac{{\omega_T}^2\omega^4}{{\omega_N}^2(\omega^2-{\omega_N}^2)^2}\left[\frac{{\omega_N}^2}{\omega^2(1-\mu^2)} - 1\right] + \frac{{\omega_L}^2}{{\omega_N}^2}\right\}$$
$$\times\left\{\frac{{\omega_N}^2}{\omega^2(1-\mu^2)} + \frac{{\omega_T}^2/2}{\omega^2-{\omega_N}^2} - 1\right\}^{-1} \qquad (15b)$$

and by substituting (RATCLIFFE, 1959, p. 53; 7)

$$X = \frac{{\omega_N}^2}{\omega^2},\ Y_L = \frac{\omega_L}{\omega},\ Y_T = \frac{\omega_T}{\omega}$$ we get the final result:

$$\mu\mu' = 1 + \frac{(1-\mu^2)^2}{2}P \qquad (16a)$$

$$P = \left\{\frac{{Y_T}^2}{X(1-X)^2}\left[\frac{X}{1-\mu^2} - 1\right] + \frac{{Y_L}^2}{X}\right\}\left\{\frac{X}{1-\mu^2} + \frac{{Y_T}^2/2}{1-X} - 1\right\}^{-1}. \qquad (16b)$$

For purposes of calculation we can rewrite a simpler form by using the following substitutions:

$$\alpha = \frac{X}{1-\mu^2} - 1 \tag{17a}$$

$$\beta = \frac{{Y_T}^2/2}{1-X} = \frac{{\omega_T}^2/2}{\omega^2 - {\omega_N}^2} \tag{17b}$$

$$\gamma = \frac{{Y_L}^2}{X} = \left(\frac{\omega_L}{\omega_N}\right)^2 \tag{17c}$$

$$\mu\mu' = 1 + \frac{(1-\mu^2)^2}{2} P \tag{17d}$$

$$P = \left[\frac{2\alpha\beta}{X(1-X)} + \gamma\right][\alpha + \beta]^{-1}. \tag{17e}$$

Expressions (16) or (17) are much simpler than those given by SHINN and WHALE (1952), MULLALY (1956), WHALE and STANLEY (1950) and BECKER (1955). In case $Y_T = Y_L = 0$ we have in (17) $\beta = \gamma = 0$ and we get $P = 0$ and $\mu\mu' = 1$ as in (10).

In case $Y_L = 0$ we have $\gamma = 0$ and we get in (17):

$$\mu\mu' = 1 + \frac{(1-\mu^2)^2}{2} \frac{1}{X(1-X)} \frac{2\alpha\beta}{\alpha+\beta} \tag{18}$$

and in case $Y_T = 0$, $\beta = 0$ and we get:

$$\mu\mu' = 1 + \frac{(1-\mu^2)^2}{2} \frac{\gamma}{\alpha}. \tag{19}$$

VOL. 20 1961

### 3. THE REFLECTION POINT $X = 1$

Near the reflection point of the ordinary wave we have $|1 - X| \ll 1$. The behaviour of $\mu^2$ there could be found from (2):

$$\mu^2 = \left\{1 - X - \frac{{Y_T}^2/2}{1-X} \pm \left[\left(\frac{{Y_T}^2/2}{1-X}\right)^2 + {Y_L}^2\right]^{1/2}\right\} \times \left\{1 - \frac{{Y_T}^2/2}{1-X} \pm \left[\left(\frac{{Y_T}^2/2}{1-X}\right)^2 + {Y_L}^2\right]^{1/2}\right\}^{-1}. \tag{20}$$

(20) may be rewritten as follows if we substitute $1 - X = \varepsilon$

$$\mu^2 = \frac{\varepsilon^2 - {Y_T}^2/2 \pm [({Y_T}^2/2)^2 + \varepsilon^2 {Y_L}^2]^{1/2}}{\varepsilon - {Y_T}^2/2 \pm [({Y_T}^2/2)^2 + \varepsilon^2 {Y_L}^2]^{1/2}}. \tag{21}$$

Since $\varepsilon$ is small we may write:

$$\left[\left(\frac{{Y_T}^2}{2}\right)^2 + \varepsilon^2 {Y_L}^2\right]^{1/2} = \frac{{Y_T}^2}{2}\left[1 + \left(\frac{Y_L}{{Y_T}^2/2}\right)^2 \varepsilon^2\right]^{1/2} = \frac{{Y_T}^2}{2}\left[1 + \frac{1}{2}\left(\frac{Y_L}{{Y_T}^2/2}\right)^2 \varepsilon^2 + \ldots\right] \tag{22}$$

Let us take only two terms in the expansion and substituting (22) in (21) we get:

$$\mu^2 = \frac{\varepsilon^2 - Y_T^2/2 \pm [Y_T^2/2 + (Y_L^2/Y_T^2)\varepsilon^2]}{\varepsilon - Y_T^2/2 \pm [Y_T^2/2 + (Y_L^2/Y_T^2)\varepsilon^2]}. \tag{23}$$

In case of the ordinary wave we take the positive sign and get in (23):

$$\mu^2 = \frac{\varepsilon^2[1 + (Y_L/Y_T)^2]}{\varepsilon[1 + (Y_L/Y_T)^2\varepsilon]}. \tag{24}$$

Since $1 - X = \varepsilon \to 0$ (24) becomes near the reflection point:

$$\mu^2 \cong \left[1 + \left(\frac{Y_L}{Y_T}\right)^2\right]\varepsilon, \quad \varepsilon = 1 - X. \tag{25}$$

(25) expresses the behaviour of $\mu^2$ near the reflection point. (25) may be rewritten as:

$$1 - \mu^2 \cong 1 - \varepsilon - \left(\frac{Y_L}{Y_T}\right)^2 \varepsilon = X - \left(\frac{Y_L}{Y_T}\right)^2 \varepsilon \tag{26}$$

or in the form:

$$1 - \mu^2 \cong X\left[1 - \frac{Y_L^2}{XY_T^2}\varepsilon\right] = X[1 - \delta] \tag{27a}$$

$$\delta = \frac{Y_L^2}{XY_T^2}\varepsilon, \quad 1 - X = \varepsilon. \tag{27b}$$

Substituting $\varepsilon = 1 - X$ and (27a) in (16) we get:

$$\mu\mu' = 1 + \frac{X^2}{2}(1 - \delta)^2\left[\frac{Y_T^2}{X\varepsilon^2}\left(\frac{1}{1 - \delta} - 1\right) + \frac{Y_L^2}{X}\right] \times \left[\frac{1}{1 - \delta} - 1 + \frac{Y_T^2/2}{\varepsilon}\right]^{-1} \tag{28}$$

and since $(1/1 - \delta) \cong 1 + \delta$, (28) may be written as:

$$\mu\mu' = 1 + \frac{X^2}{2}(1 - \delta)^2\left(\frac{Y_T^2}{X}\frac{\delta}{\varepsilon^2} + \frac{Y_L^2}{X}\right)\left(\delta + \frac{Y_T^2/2}{\varepsilon}\right)^{-1} \tag{29}$$

and (29) may be rewritten as:

$$\mu\mu' = 1 + \frac{X}{2}(1 - \delta)^2\left[\frac{Y_T^2(\delta/\varepsilon) + Y_L^2\varepsilon}{\delta\varepsilon + (Y_T^2/2)}\right]. \tag{30}$$

Neglecting values of second order we get by using (27b):

$$\mu\mu' = 1 + \frac{X}{2}(1 - \delta)^2\,2\frac{\delta}{\varepsilon} = 1 + (1 - \delta)^2\left(\frac{Y_L}{Y_T}\right)^2 \cong 1 + \left(\frac{Y_L}{Y_T}\right)^2. \tag{31}$$

From (25) and (31) we see that near the reflection point $X = 1$ the ordinary wave

refraction index $\mu$ and group refraction index $\mu'$ behave as:

$$\mu \cong \left[1 + \left(\frac{Y_L}{Y_T}\right)^2\right]^{1/2} (1 - X)^{1/2} \tag{32a}$$

$$\mu' \cong \left[1 + \left(\frac{Y_L}{Y_T}\right)^2\right]^{1/2} (1 - X)^{-1/2}. \tag{32b}$$

(32) is very suitable for the calculation of $\mu$ and $\mu'$ near the reflection point, which was the main problem in the previous methods. In case of $Y_L = Y_T = 0$ we get near the reflection point from (10):

$$\mu\mu' = 1, \qquad \mu = (1 - X)^{1/2}, \qquad \mu' = (1 - X)^{-1/2}. \tag{33}$$

In case $Y_L = 0$, $Y_T \neq 0$ (ordinary wave) we get:

$$\mu\mu' = 1, \qquad \mu = (1 - X)^{1/2}, \qquad \mu' = (1 - X)^{-1/2}. \tag{34}$$

In case $Y_T = 0$, $Y_L \neq 0$ (ordinary wave) we do not get reflection at $X = 1$ at all.

## REFERENCES


| | | |
|---|---|---|
| BECKER W. | 1955 | *Arch. Elektr. Übertr.* **9**, 277. |
| MULLALY R. F. | 1956 | *J. Atmosph. Terr. Phys.* **9**, 322. |
| RAMO S. and WHINNERY J. R. | 1953 | *Fields and Waves in Modern Radio*, p. 48. Wiley, New York. |
| RATCLIFFE J. A. | 1959 | *The Magneto-Ionic Theory and its Applications to the Ionosphere*. Cambridge University Press. |
| SHINN O. H. and WHALE H. A. | 1952 | *J. Atmosph. Terr. Phys.* **2**, 85. |
| STRATTON J. A. | 1941 | *Electromagnetic Theory*, pp. 330–340. McGraw-Hill, New York. |
| WHALE H. A. and STANLEY J. P. | 1950 | *J. Atmosph. Terr. Phys.* **1**, 82. |




VOL. 20 1961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 195 to 199. Pergamon Press Ltd. Printed in Northern Ireland

# Study of variations of the forbidden oxygen lines in the nightglow

M. C. Pande and S. S. Varma
Uttar Pradesh State Observatory, Naini Tal, India

(*Received* 9 *September* 1960)

**Abstract**—The variations of the forbidden oxygen lines in the nightglow are investigated photoelectrically using a fixed orientation photometer set in a star free region near the north celestial pole. For the red lines our observations show that the slope of the evening twilight enhancement portion of the diurnal curve indicates a possible change with season; for the 5577 Å emission, the occurrence of a maximum around midnight is confirmed. Little correlation of these emissions with solar activity could be established.

## Introduction

In the visible region of airglow emission spectra the prominent atomic lines are 5577 Å ($^1S \rightarrow {}^1D$), 6300–6364 Å ($^1D \rightarrow {}^3P$) and the *D*-lines of sodium. The diurnal and seasonal variations of forbidden oxygen lines in the green and red regions of the spectrum have been studied by Roach *et al.* (1953), Dufay and Tcheng Mao-lin (1946), Barbier (1959) and others. For 5577 Å a midnight maximum, first indicated by Dufay and Tcheng Mao-lin (1946) from spectrograms of moderate dispersion, has been established by Roach and Pettit (1951) by photoelectric techniques. For 6300–6364 Å Elvey and Farnsworth (1942) found a considerable evening twilight enhancement with a more or less steady level of radiation during the night and a slight enhancement towards morning hours. The "twilight flash" was first indicated by Slipher (1933).

The present investigation has been carried out at the Uttar Pradesh State Observatory, Naini Tal (latitude 29·5° N, longitude 79·5° E, height above sea level 6900 ft) with a view to examine further the nature of variations of these radiations.

## Instrumental Technique

We have used a fixed orientation photometer for the purpose. Much useful information can be obtained by this technique. The greatest advantage of the method is that the background due to the unresolved faint stars does not change much and hence corrections for the illumination contribution from the varying galactic background would be unnecessary. The fixed orientation photometer also gives a higher time resolution for examining a single region of the sky in one radiation. The disadvantage of the fixed orientation photometer is that it misses the emission patches which can be picked up by the scanning photometers in all-sky surveys.

The observations reported herein were made photoelectrically using a 931-A photomultiplier tube and interference filters. The photometer had an altazimuth mounting and, being of the fixed orientation type, it was easy to set it on a region near the north celestial pole devoid of bright stars. An objective lens of 3 in. aperture and 17 in. focal length focussed an area of the sky onto a diaphragm of 7·5 mm diameter arranged to lie in the focal plane. The airglow radiations from an area of the sky equal to 0·80 square degrees thus fell on the photocathode after

passing through a Fabry lens and an interference filter. We used interference filters manufactured by Baird Associates for isolating the emission lines and the continuum. These filters have a width at half intensity of about 55 Å and a peak transmission of 70 per cent and isolate the regions centred on 5300, 5577 and 6332 Å. Both 5577 Å and 6300–6364 Å, particularly the latter, are situated very near to the strong OH bands. The two bands are 5593 Å (intensity one-sixteenth of 5577 Å) and at 6257 Å (intensity almost equal to that of the red forbidden oxygen lines) (Roach, 1957). From the filter transmission characteristics we find that the OH bands contribute 0·044 of the radiation at peak transmitted by the 5577 Å filter and an insignificant amount to radiation received by the photomultiplier through the red filter. We tilted the filters so that the transmission peaks would be centred to within 2 Å of the wavelengths of interest. This was done with the aid of sunlight and a medium dispersion monochromator available in our laboratory. The output of the photomultiplier was fed to a d.c. amplifier having highly linear characteristics and coupled to a 0–1 mA Esterline Angus recorder. A button of radium paint in the photometric apparatus was periodically used to check the overall stability of the electronics.

Observations were made on moonless nights and when the transparency of the sky was extremely stable. Photoelectric work carried out at this Observatory simultaneously on these nights indicated good stability of the values of atmospheric extinction on these nights. In general the altitude at which these observations were carried out ensured low extinction values with good stability.

VOL. 20 1961

## Nature of Variations

For both 5577 Å and 6300 Å the nights can be isolated into two groups, one showing smooth variation and another characterized by small irregular fluctuations superposed on the general smooth variation. The variations of the ratio of intensity at 5577 Å to the intensity at 5300 Å expressed in magnitudes [$-2{\cdot}5 \log (I\ 5577/I\ 5300)$] are shown in Fig. 1. The variations on different nights are not identical. The ratio attains a maximum within 3 hr of midnight. This agrees well with data of Roach and Pettit (1951). The nights for which 5577 Å depicts smooth variation, show double maxima, one more prominent than the other. The shift of time of occurrence of maxima from midnight for various dates is given in Table 1. We see that the maxima occur before midnight from the middle of October to the beginning of March and after midnight during March and April. We have compared it with similar data obtained by Roach *et al.* (1953) and find that in general during the period in which those observations were made the maxima occur after midnight between the middle of October and the beginning of March. We cannot emphasize this difference much since their study covers a much longer period than ours. The red lines 6300–6364 Å show considerable enhancement in evening twilight. The morning twilight enhancement also exists, but the relative change is comparatively less, which is explained as being caused by an unequal duration of illumination by sunlight of the upper atmosphere in the two cases (Fig. 2). In the night time the ratio $I$ 6300/$I$ 5300 either stays constant or shows a minimum near midnight. Our observations show that the slope of the evening twilight enhancement portion of the curve indicates a possible change with season. The

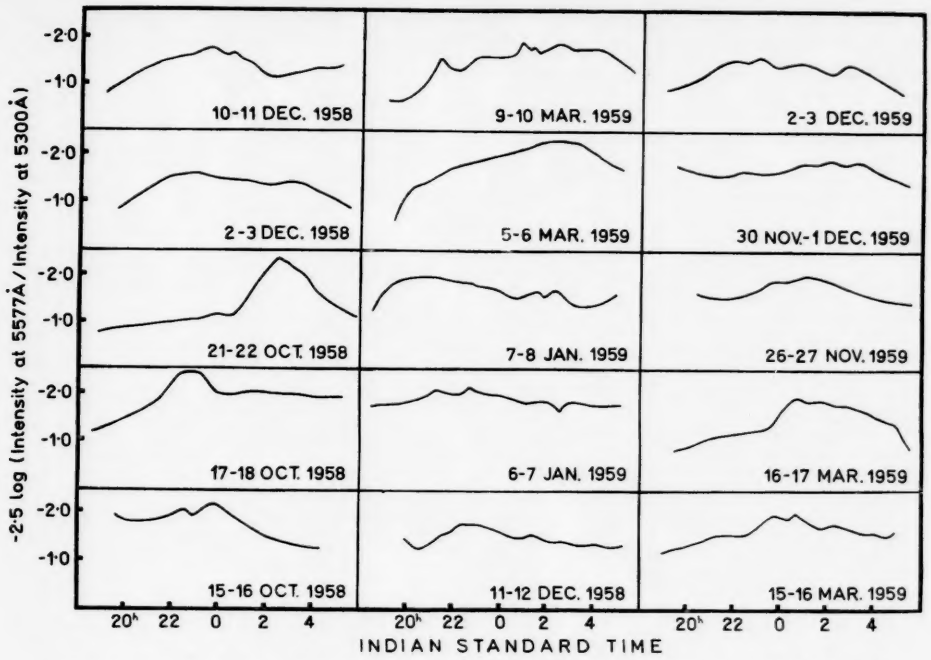


Fig. 1. Variation of the intensity of 5577 Å radiation in the night airglow at Naini Tal.

Table 1.

| Date | Nature of solar activity | Intensity of 5577 Å | Shift* of time of maximum of 5577 Å radiation from mid-night (hr) |
|---|---|---|---|
| 15/16.x.58 | Normal | Normal | +2·5 |
| 17/18.x.58 | Active | Abnormal | −1·5 |
| 22/23.x.58 | Active (SWI) | Abnormal | −1·5 |
| 2/3.xii.58 | Active (alert) | Normal | −2·3 |
| 7/8.xii.58 | Normal | Normal | −2·0 |
| 10/11.xii.58 | Active (alert) | Normal | −0·5 |
| 11/12.xii.58 | Active (alert) | Normal | −1·3 |
| 6/7.i.59 | Normal | Normal | −3·0 |
| 7/8.i.59 | Normal | Normal | −2·5 |
| 5/6.iii.59 | Normal | Normal | +2·5 |
| 10/11.iii.59 | Normal | Normal | +0·75 |
| 16/17.iii.59 | Normal | Normal | +0·75 |
| 17/18.iii.59 | Normal | Normal | +0·75 |
| 31.xi/1.xii.59 | Normal | Normal | +2·00 |
| 2/3.xii.59 | Normal | Normal | −1·00 |

* (+) indicates a shift towards later hours.

slopes are given in Table 2. The change in slope of the twilight enhancement portion of the curve is significant and it may help in an understanding of the enhancement mechanism. The change may be due to any of the two variable factors, the concentration of the reactants or more likely the rate of removal of

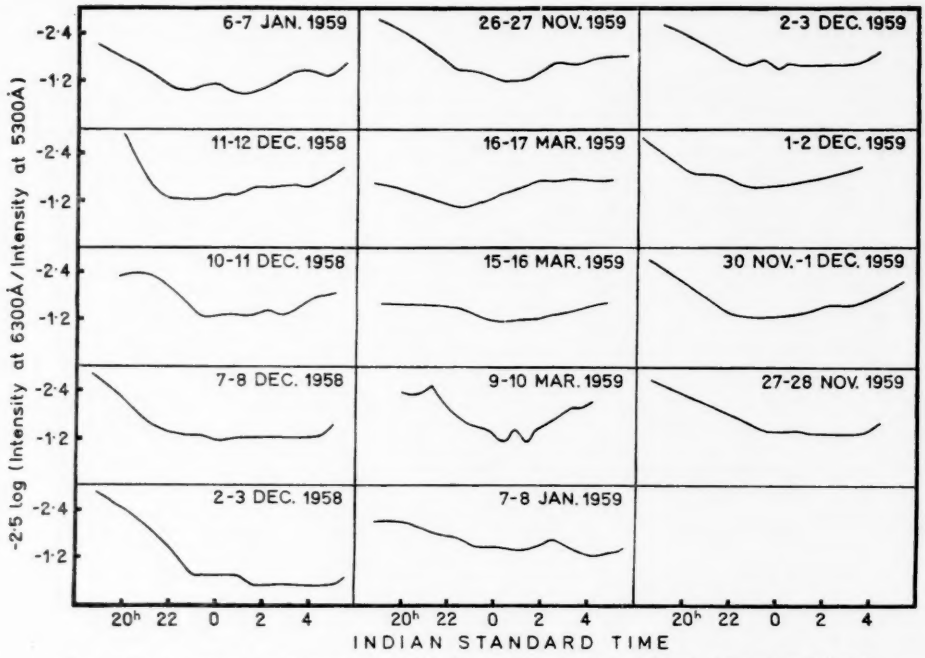


Fig. 2. Variation of the intensity of 6300 Å radiation in the night airglow at Naini Tal.

Table 2. Slope of the evening twilight enhancement potion of the diurnal variation curve for red lines

| Date | Slope | Date | Slope |
|---|---|---|---|
| 2/3.xii.58 | 1·38 | 15/16.iii.59 | 0·23 |
| 7/8.xii.58 | 1·46 | 16/17.iii.59 | 0·17 |
| 10/11.xii.58 | 1·20 | 26/27.xi.59 | 1·24 |
| 11/12.xii.58 | 4·53 | 27/28.xi.59 | 0·83 |
| 6/7.i.59 | 0·58 | 1/2.xii.59 | 1·55 |
| 7/8.i.59 | 0·46 | 2/3.xii.59 | 0·84 |

solar ultra-violet. Both these factors may vary with the season in such a manner that the evening twilight enhancement portion of the curve for summer is comparatively flatter than that during winter. The morning twilight enhancement portion during summer has a larger gradient. This point, however, needs further study.

## Effect of Solar Activity

Since low latitude aurorae are expected to occur in periods of increased solar activity (as it was during 1957–1958), 5577 Å and 6300 Å may indicate an increase

in intensity during such times. Our observations do not indicate any such correlation with solar activity. Both high and low intensities of 5577 Å are found on nights when solar activity is high. The red forbidden oxygen lines also do not show any correlation with the activity of the sun. The data, however, are too meagre to enable us to make a definite statement. The nature of the sun's activity corresponding to each night is indicated in Table 1. The criteria for the distinction are the A.G.I. warnings which were received at this Observatory during the duration of the International Geophysical Year.

*Acknowledgements*—The authors thank Dr. M. K. VAINU BAPPU and Dr. S. D. SINVHAL for their helpful suggestions. One of us (M. C. P.) is indebted to the Council of Scientific and Industrial Research, India and the other (S. S. V.) to the Scientific Research Committee, Uttar Pradesh, for financial support.

## REFERENCES


| | | |
|---|---|---|
| BARBIER D. | 1959 | *Ann. Géophys.* **15,** 412. |
| DUFAY J. and TCHENG MAO-LIN | 1946 | *Ann. Géophys.* **2,** 189. |
| ELVEY C. T. and FARNSWORTH A. H. | 1942 | *Astrophys. J.* **96,** 451. |
| ROACH F. E. | 1957 | *Ann. I.G.Y.* **4,** 118. |
| ROACH F. E. and PETTIT H. B. | 1951 | *J. Geophys. Res.* **56,** 325. |
| ROACH F. E., PETTIT H. B., WILLIAMS D. R., ST. AMAND P. and DAVIS D. N. | 1953 | *Ann. Astrophys.* **16,** 185. |
| SLIPHER V. M. | 1933 | *Trans. Amer. Geophys. Un.* (1933), 135. |


VOL.
20
961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 200 to 205. Pergamon Press Ltd. Printed in Northern Ireland

# Behaviour of broadcast frequency waves at oblique incidence during an annular eclipse

R. K. MacCrone* and F. R. N. Nabarro
Physics Department, University of the Witwatersrand

(*Received* 12 *August* 1960; *in revised form* 11 *October* 1960)

**Abstract**—Measurements taken during a solar eclipse show discontinuities in the absorption of obliquely incident waves in the broadcast band. The results are consistent with Appleton's explanation in terms of an abrupt change in the height of the point of reflection. Some evidence for a corpuscular eclipse 2 hr before the peak of the optical eclipse is also presented.

## 1. Introduction

An annular eclipse of the sun took place in South Africa on 25 December 1954. Recordings were made of the outputs of communications receivers receiving continuous waves of various frequencies transmitted from places in South Africa. The receivers were situated at Beaufort West (32°18′ S, 22°36′ E), which lay in the belt of maximum phase which occurred at approximately 0800 hours S.A.S.T. (S.A.S.T. = U.T. + 2 hr). The total duration of the eclipse was in all cases about 2 hr.

VOL. 20 1961

Table 1. Conditions of propagation

| Transmitter | Frequency (kc/s) | Reflection point | | $\phi$ |
|---|---|---|---|---|
| | | Lat. S | Long. E | |
| Port Elizabeth | 1313 | 33°9′ | 24°7′ | 70°0′ |
| Cape Town | 728 | 33°7′ | 20°31′ | 82°30′ |
| Grahamstown | 620 | 32°49′ | 24°34′ | 77°0′ |
| Cape Town | 557 | 33°7′ | 20°31′ | 82°30′ |
| Lourenço Marques (Receiver at Johannesburg) | 917 | 24°35′ | 30°41′ | 82°30′ |

Table 1 shows the transmitters used, their frequencies, the geographical co-ordinates of the points of reflection and the angle $\phi$ between the horizontal direction of propagation and the total magnetic field of the earth at the point of reflection. The total magnetic field was about 0·31 gauss. For supplementary observations with the transmitter at Lourenço Marques and the receiver at Johannesburg, the total field was about 0·30 gauss.

Recordings were also made 2 days before the eclipse and 1 day after to provide control results with which to compare the results observed on the day of the eclipse.

* Now at the School of Metallurgical Engineering, University of Pennsylvania.

## 2. Optical Eclipse

Figs. 1, 2, 3 and 4 show the behaviour of the 1313, 728, 620 and 557 kc/s transmissions respectively. The smoothed average of the control days is simultaneously shown for comparison.

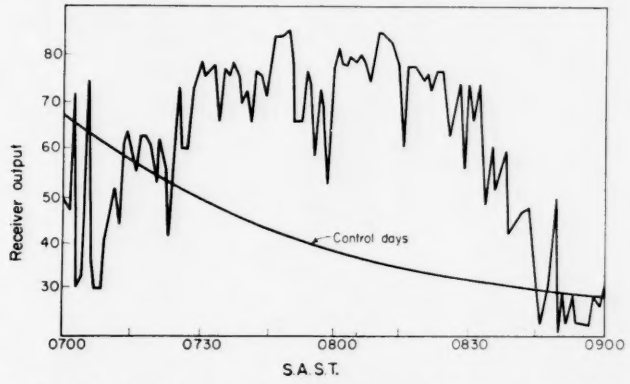


Fig. 1. Received signal as a function of time over period of annular eclipse, with smoothed average of control days for comparison. Oblique incidence. 1313 kc/s.

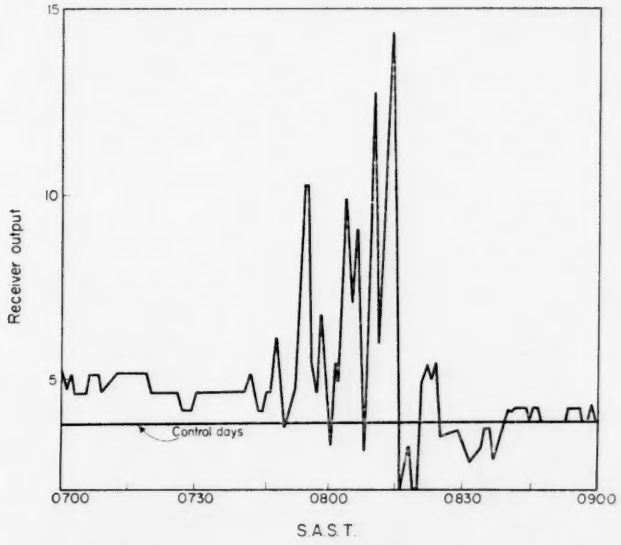


Fig. 2. Received signal as a function of time over period of annular eclipse, with smoothed average of control days for comparison. Oblique incidence. 728 kc/s.

The behaviour indicated in Fig. 1, 1313 kc/s, differs from the behaviour shown in Figs. 2 and 3 for 728 and 620 kc/s respectively. In Fig. 1 the amplitude of the received sky wave increased gradually above control during the $\frac{1}{2}$ hr preceding the maximum phase of the eclipse and then decreased gradually during the $\frac{3}{4}$ hr following the maximum phase. It should be noticed that the receiver became

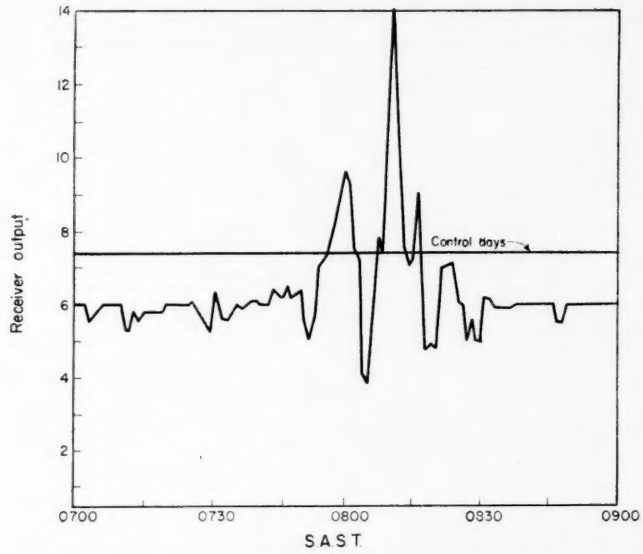


Fig. 3. Received signal as a function of time over period of annular eclipse, with smoothed average of control days for comparison. Oblique incidence. 620 kc/s.

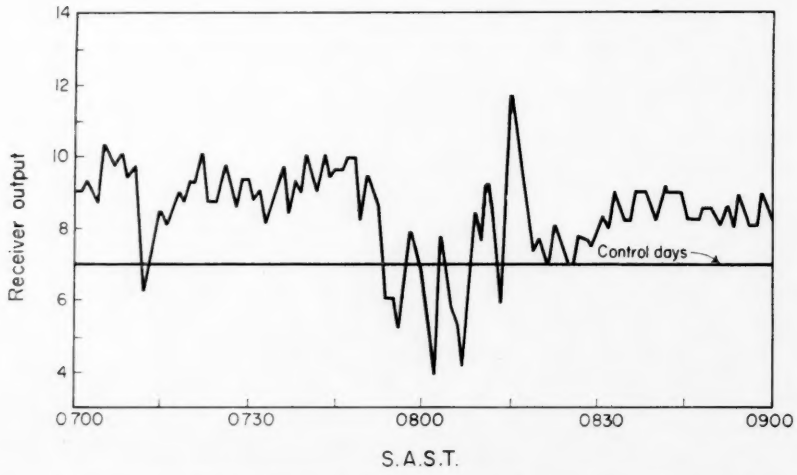


Fig. 4. Received signal as a function of time over period of annular eclipse, with smoothed average of control days for comparison. Oblique incidence. 557 kc/s.

saturated at the highest intensities. In Figs. 2, 3 and 4, however, the sky wave amplitudes increased appreciably above control rather abruptly in a period of less than $\frac{1}{4}$ hr before the maximum phase, and decreased again in the 25 min following the maximum phase. In Fig. 4, 557 kc/s, the effects are smaller, but confined to the same time interval.

We propose the following explanation, which is essentially the same as that



VOL. 20 1961

given by APPLETON (1928) for the very similar, though more detailed, observations made during the eclipse of 29 June 1927:

It is assumed, after APPLETON and PIGGOTT (1954), that the absorption in the *D*-region is greater than the absorption in the *E*-region. During the period 0700–0900 hours S.A.S.T. on normal days, no appreciable sky wave amplitudes were received on the three lower frequencies, which means that the absorption in the *D*-region was very high. The reflection points for these frequencies are in the *D*-region, according to the model of FEJER (1955). $N_e$ for reflection at these frequencies and obliquities is about $1{\cdot}5 \times 10^3$ electrons cm$^{-3}$. During the eclipse,

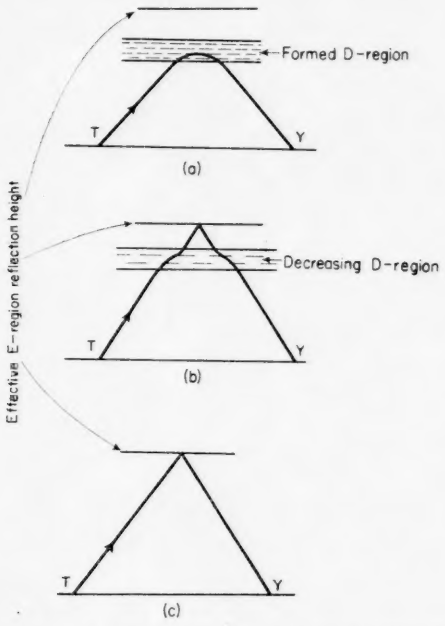


Fig. 5. Change of height of reflection during increasing phase of eclipse. (a) Reflection in *D*-region (b) Maximum electron density in *D*-region has fallen just below critical value (c) Negligible electron density in *D*-region.

the decreasing electron density in the *D*-region results in an abrupt change of the reflection point from the *D*- to the *E*-region, as indicated in Fig. 5, with a resulting abrupt decrease in absorption.

This behaviour is not shown at 1313 kc/s, since reflection is considered to take place from the *E*-region at all times, and the decreasing absorption as the maximum phase is approached is indicative of the decreasing electron density in the *D*-region during this time.

As additional evidence for the hypothesis of an abrupt change of the reflection point at oblique incidence, the results of C.W. recordings at Johannesburg of transmissions from Lourenço Marques at 917 kc/s during night–day transition periods in March 1956 are shown in Fig. 6. The vertical bars show the scatter caused by ground wave interference, and data are presented for 2 days to show

the reproducibility of the observations. During the formation between 0545 and 0600 hours S.A.S.T. of the *D*-region, which is absent at night, there is an abrupt increase in the absorption, indicative of the process considered.

The abrupt increase of intensity of oblique waves just before the time of maximum phase has also been observed by Stoffregen (1956) and others. It has not, so far as the authors are aware, been observed at vertical incidence (see Piggott, 1956). There may be alternatives to the explanation proposed here, but it seems clear that, whatever the mechanism, small changes in electron density result in large changes in received amplitude.

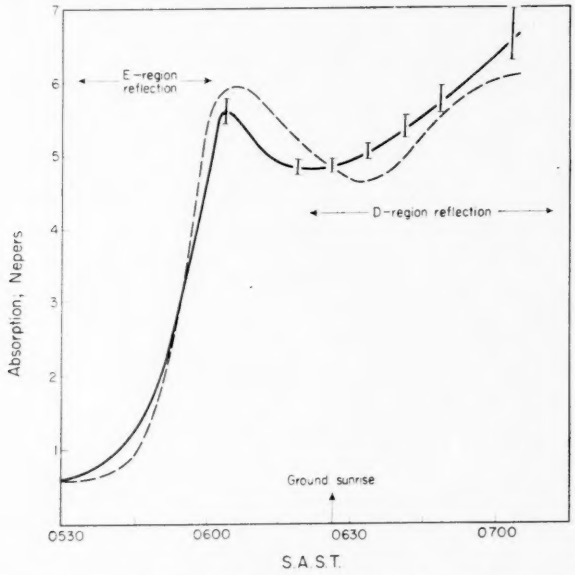


Fig. 6. Normal day. Variation of signal received in Johannesburg from transmission at Lourenço Marques over period of sunrise. Upright bars, scatter attributed to ground wave interference. Dotted curve, a second day, showing that the principal features are reproducible. 917 kc/s.

## 3. Particle Eclipse

Chapman (1932) has considered the effect of the flux of solar particles on the ionosphere. Assuming the speed of the particles to be 1600 km $sec^{-1}$, he showed that a corpuscular eclipse should occur about 1000 miles to the east of the optical eclipse, that the maximum phase of the corpuscular eclipse should occur about 2 hr before the maximum phase of the optical eclipse, that the area of the earth experiencing a corpuscular eclipse would be very much larger than that experiencing an optical eclipse and that the corpuscular eclipse would last about 24 min. Particle eclipse phenomena have previously been observed by Sen Gupta and Mitra (1954).

Since the eclipse observed at Beaufort West had a roughly East–West track, and at the expected time of the corpuscular eclipse the reflection point was moving rapidly from the *E*- to the *D*-layer, the conditions were particularly suitable for observing the corpuscular eclipse.



VOL. 20 1961

The behaviour of the 728, 620 and 557 kc/s transmissions on the day of the eclipse at the expected time of the particle eclipse does differ from the behaviour shown on the control days, and we attribute the difference to the occurrence of a particle eclipse. Due, however, to the inherent scatter in the results obtained by this method of observation, the difference may not be significant.

We also observed, as did APPLETON (1928), anomalies on the morning following the eclipse. On normal days the fall from night-time to day-time intensity is interrupted by a transient rise. On the day following the eclipse this rise was not observed. Again the effect is barely significant, but since we had noticed it before we were aware of APPLETON's similar observation we believe that it is real.

*Acknowledgements*—We would like to thank the South African Broadcasting Corporation and the Radio Clube de Moçambique for providing transmissions during the awkward times we required, and to thank Dr. J. A. FEJER for his advice and valuable criticism. We would also like to thank A.E.C.I. Ltd. for a substantial grant to one of us during the time this work was being carried out. Equipment was kindly lent by the following South African Laboratories: Bernard Price Institute for Geophysical Research; G.P.O. Research Laboratory; National Physical Research Laboratory; Telecommunications Research Laboratory. Accommodation was provided by the Department of Transport and the Municipality of Beaufort West. In addition to the authors, the following students of the University of the Witwatersrand assisted in the tedious observations: L. BURGER, R. D. ELLIOT, D. B. A. EPSTEIN, R. HOLLOWAY, K. ISEBECK, E. LEADER, J. W. MATTHEWS, M. D. NABARRO, E. A. NEWTON, J. SOSNOVIC.

OL.
20
961

## REFERENCES


| | | |
|---|---|---|
| APPLETON E. V. | 1928 | *J. Instn Elect. Engrs* **66,** 872 |
| APPLETON E. V. and PIGGOTT W. R. | 1954 | *J. Atmosph. Terr. Phys.* **5,** 151. |
| CHAPMAN S. | 1932 | *Mon. Not. R. Astr. Soc.* **92,** 413. |
| FEJER J. A. | 1955 | *J. Atmosph. Terr. Phys.* **7,** 322. |
| PIGGOTT W. R. | 1956 | *Solar Eclipses and the Ionosphere*, p. 106. Pergamon Press, London. |
| SEN GUPTA P. K. and MITRA S. N. | 1954 | *Nature, Lond.* **173,** 814. |
| STOFFREGEN, W. | 1956 | *Solar Eclipses and the Ionosphere*, p. 57. Pergamon Press, London. |

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 206 to 219. Pergamon Press Ltd. Printed in Northern Ireland

# RESEARCH NOTES

## Wave-guide mode propagation of very low frequency radio waves

(*Received* 9 *November* 1960)

RECENTLY there has been some interest in the non-reciprocity existing when very low frequency radio waves are propagated between widely separated points. Thus TREMELLEN (1950) stated that over long distance routes of 6000 km or more the signal intensity was considerably less when receiving from a transmitter situated to the east than from one to the west; CROMBIE (1958) has shown that there is considerable evidence supporting these observations, which were originally made by ROUND *et al.* (1925). BARBER and CROMBIE (1959) calculated the reflection coefficient of a sharply bounded homogeneous ionosphere with a superimposed horizontal magnetic field perpendicular to the plane of propagation, and found it to be numerically greater for waves at nearly grazing incidence from the west than for those at similar incidence from the east. CROMBIE (1960) has used these results in a wave-guide mode theory, showing that for the least attenuated mode the attenuation is less for W–E than for E–W propagation. The theory was given for a sharply bounded homogeneous ionosphere with a horizontal transverse magnetic field, and thus applied only to propagation within the plane of the equator.

The purpose of this note is to report some preliminary calculations in which a more realistic model of the ionosphere is used. The electron density increases exponentially with height and the magnetic field is inclined at 30° to the vertical in a plane perpendicular to the plane of propagation. Thus the results apply to W–E and E–W propagation in northern temperate latitudes, and use a model of the ionosphere which is probably closer to the true profile than is a sharply bounded homogeneous medium.

The field in a wave-guide mode can be considered as two crossing waves whose normals make complex angles $\pm\theta$ with the horizontal. The attenuation of the mode is proportional to the imaginary part of $\cos\theta$. The values of $\theta$ for the various modes were calculated by a method outlined by BUDDEN (1957) which uses a matrix admittance variable **A**. (In BUDDEN's report **V** is written for **A**.) This must satisfy a differential equation which was given by BARRON and BUDDEN (1959). For any given $\theta$ the differential equation can be solved by numerical integration to give the value of **A** at some level in the lowest part of the ionosphere, which will be called the base of the ionosphere. From this the reflection coefficient matrix **R** at this level can be found, and thence its value $\mathbf{R}_0$ at the ground, which is given by:

$$\mathbf{R}_0 = \mathbf{R} \,.\, \exp(-2ihk \sin\theta) \tag{1}$$

where $h$ is the height of the base of the ionosphere. The only acceptable values of $\theta$ are those for which the boundary conditions at the ground are satisfied, and these require that:

$$\det\left\{\mathbf{R}_0 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right\} = 1. \tag{2}$$

BUDDEN's method consists in expanding $\mathbf{A}_1$, the value of **A** at the base of the ionosphere, as a Taylor series in $\sin\theta$. This contains only even powers of $\sin\theta$ the coefficients of which were calculated for the first five non-zero terms, that is up to the eighth derivative of **A**.



VOL. 20 1961

By summing the series **A** could be found for any given $\theta$; the series was found to converge rapidly enough provided that $|\sin\theta| < 0{\cdot}7$. $\mathbf{R}_0$ was then computed from (1) and its value used repeatedly in (2) together with a process of successive approximation until a value of $\sin\theta$ was found which was consistent with equation (2).

The results shown are based on the following details using the standard notation. The earth is assumed to be flat and the ground perfectly conducting, the constant collision frequency is given by $\nu = 3 \times 10^6\ \text{sec}^{-1}$, the gyro-frequency by 1·28 Mc/s. The variation in electron density is defined by $X = \exp(\alpha z)$, where $z$ is measured vertically upwards from

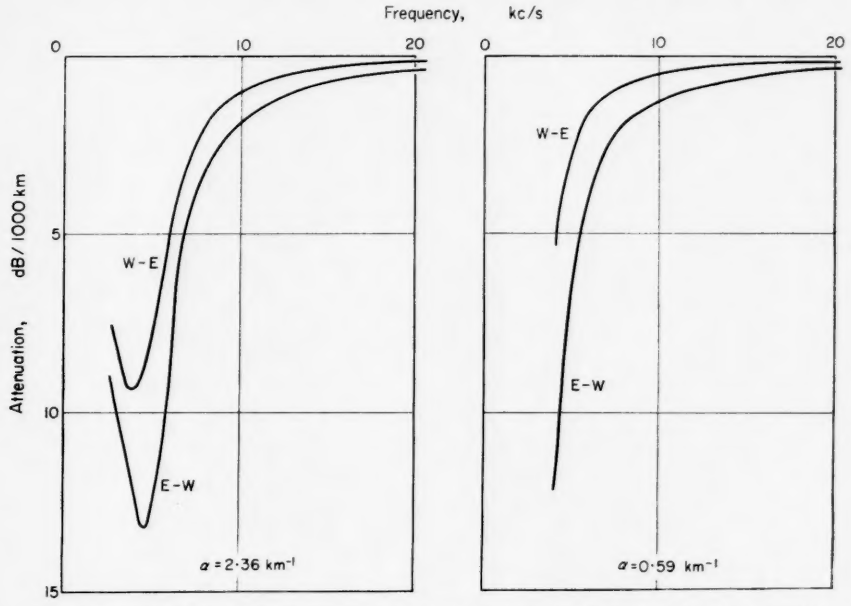


Fig. 1. Curves for the least attenuated wave guide mode. Electron density $N$ in ionosphere proportional to $\exp(\alpha z)$. At $h_1 = 80$ km, $N = 3{\cdot}175$. Flat earth. Constant collision frequency $= 3 \times 10^6\ \text{sec}^{-1}$. Earth's magnetic field at 30° to the vertical.

a level near the base of the ionosphere. The quantity $h_1$ is the distance between the ground and the level where $N = 3{\cdot}175$; this corresponds with the level where $X$ is unity for 16 kc/s. Fig. 1 shows two pairs of curves relating attenuation to frequency which refer to a very "sharp" ionosphere defined by $\alpha = 2{\cdot}36\ \text{km}^{-1}$, and a more realistic gradual model defined by $\alpha = 0{\cdot}59\ \text{km}^{-1}$. In both cases the upper curve is the less attenuated and refers to W–E propagation, the lower one to the opposite direction. The results apply to quasi-TM modes of least attenuation. They were computed for several values of $h_1$, but those shown only refer to 80 km. In Fig. 2 curves are given which relate $h_1$ to the attenuation for four frequencies and the four least attenuated modes. These show that as $h_1$ decreases the attenuation of the mode increases and the non-reciprocity becomes more apparent. It is interesting to note that the attenuation curves for a vertically polarized wave incident upon similar models of the ionosphere, but without an applied field, are very close to those given for the E–W case. These conclusions appear to be in accord with those of other authors mentioned above.

H. G. Martin

*University Mathematical Laboratory*
*Cambridge*

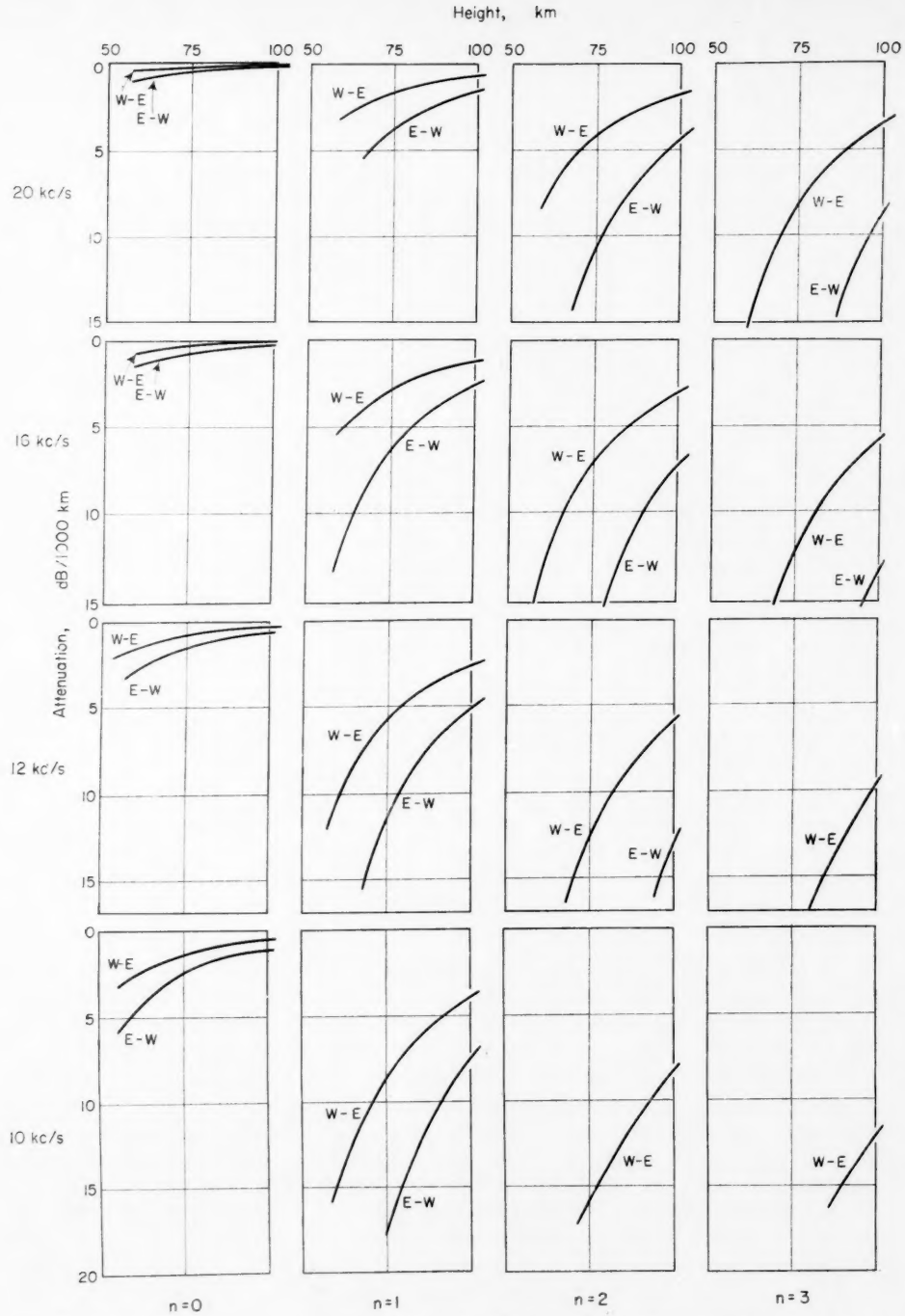


Fig. 2. Attenuation curves for wave guide modes. Electron density $N$ in ionosphere proportional to exp ($\alpha z$) where $\alpha = 2{\cdot}36$ km$^{-1}$. The abscissae are heights where $N = 3{\cdot}175$. Flat earth. Earth's magnetic field at 30° to the vertical. Constant collision frequency $= 3 \times 10^6$ sec$^{-1}$.



VOL. 20 1961

REFERENCES

| | | |
|---|---|---|
| BARBER N. F. and CROMBIE D. D. | 1959 | *J. Atmosph. Terr. Phys.* **16,** 37. |
| BARRON D. W. and BUDDEN K. G. | 1959 | *Proc. Roy. Soc.* A **249,** 387. |
| BUDDEN K. G. | 1957 | *Rep. Nat. Bur. Stand.* **5050.** |
| CROMBIE D. D. | 1958 | *J. Atmosph. Terr. Phys.* **12,** 110. |
| CROMBIE D. D. | 1960 | *J. Res. Nat. Bur. Stand.* D **64,** 265. |
| ROUND H. J., ECKERSLEY T. L., TREMELLEN K. W. and LUNNON F. C. | 1925 | *J. Instn Elect. Engrs* **63,** 933. |
| TREMELLEN K. W. | 1950 | *Marconi Rev.* **13,** 153. |


## The analysis of night-time $h'(f)$ records

(*Received* 10 *November* 1960)

THE REAL HEIGHTS of the ionospheric layers are normally determined by analysing the ordinary ray trace of sweep-frequency virtual height records. These $h'(f)$ records only extend down to some limiting frequency $f_{\min}$, and it is usually assumed that the virtual height at lower frequencies is constant and equal to the value $h'_{\min}$ at $f_{\min}$. This assumption

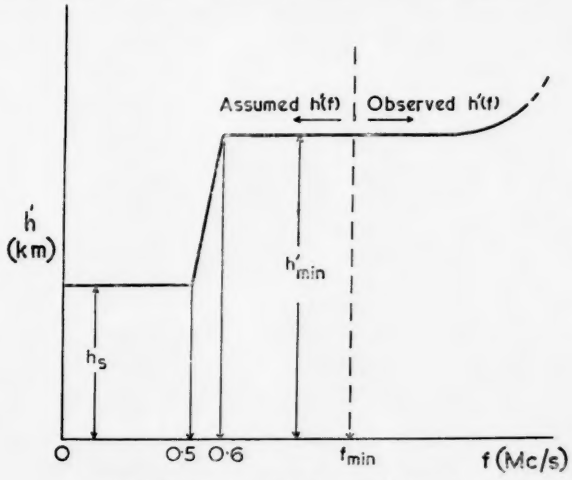


Fig. 1. The assumed $h'(f)$ curve below $f_{\min}$.

causes the calculated real heights to be greater than the true values. The error is largest at night, when the calculated heights of the $F$-layer are commonly about 20 km too high at a plasma frequency ($f_N$) of 2 Mc/s, and 7 km high at 5 Mc/s (TITHERIDGE, 1959a, b).

This error can be considerably reduced if the virtual height is assumed to be equal to $h'_{\min}$ from $f_{\min}$ down to a frequency just greater than 0·5 Mc/s (say 0·6 Mc/s) but is assumed to be equal to a value $h_s$ for frequencies from 0 to 0·5 Mc/s so that the form of the complete ordinary ray $h'(f)$ curve to be analysed is as shown in Fig. 1. The height $h_s$ is chosen in such a way that $h'_{\min} - h_s$ is proportional to the amount of low-lying ionization and the analysis of the complete $h'(f)$ curve illustrated in Fig. 1 will provide an $N(h)$ profile which is a good

approximation to the profile which would have been deduced if the $h'(f)$ curve had been observed down to zero frequency. This paper describes how the starting height $h_s$ may be determined.

The amount of this underlying ionization is approximately proportional to $h_x' - h'_{\min}$, where $h_x'$ is the observed virtual height of the extraordinary ray, of frequency $f_x$, which is reflected at the same true height as the ordinary ray of frequency $f_{\min}$ (TITHERIDGE, 1959a, Section 2). Since the amount of underlying ionization in the calculated real height curve is proportional to the assumed value of $h'_{\min} - h_s$, we require that $h'_{\min} - h_s = K(h_x' - h'_{\min})$. The value of $K$ depends on the value of $f_{\min}$ and on the strength and direction of the

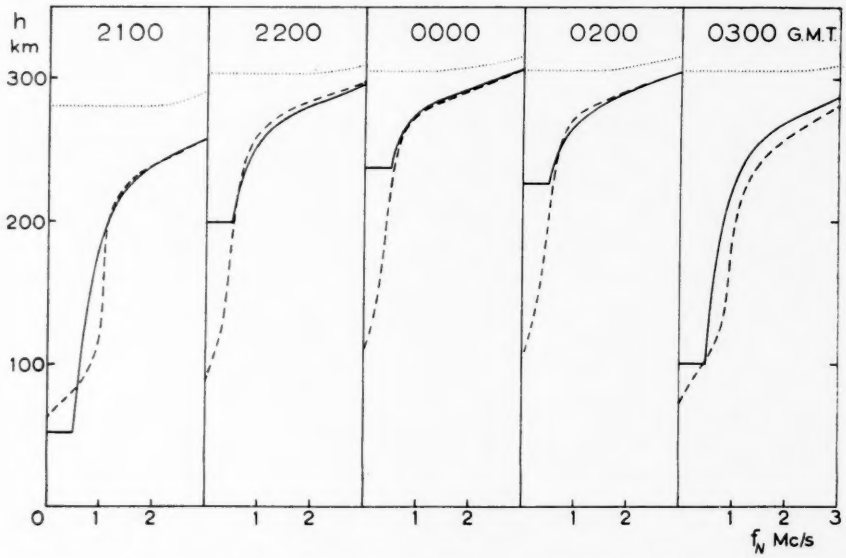


Fig. 2. Calculated heights at Slough, June 1957. Each curve gives the mean result from three of the International Quiet Days.

magnetic field. It can be determined empirically, for any desired conditions, by calculating the value of $h_x'$ resulting from a real height curve with known values of $h'_{\min}$ and $h_s$. Thus for conditions in S.E. England we get $h'_{\min} - h_s = 4{\cdot}6(h_x' - h'_{\min})$ when $f_{\min}$ is 1·4 times the gyrofrequency $f_H$.

Some typical results are shown in Fig. 2. The dotted lines give the heights calculated by the normal methods in which $h'$ is assumed to be constant and equal to $h'_{\min}$ at frequencies below $f_{\min}$. The solid lines give the result of starting the calculations at 0·5 Mc/s with the starting height $h_s = h'_{\min} - 4{\cdot}6\ (h_x' - h'_{\min})$. The calculated heights then agree exactly with the measured virtual heights of the ordinary ray, and also agree with the virtual height of the extraordinary ray at the frequency $f_x = 2f_H$. The broken line curves are taken from TITHERIDGE (1959a) and were calculated to agree with the virtual heights of both the ordinary and extraordinary rays, at all frequencies, to within the experimental error of about 1 km. The solid line curves are seen to agree quite well with these more accurate results at frequencies above $f_{\min}$ (1·5 Mc/s).

This agreement is maintained at all seasons and sunspot numbers. 120 night-time $h'(f)$ records taken at Slough were analysed by each of the three methods used in Fig. 2. The normal analysis neglecting the effect of the underlying ionization differed from the results



VOL. 20 1961

of the full analysis by a mean amount of 21 km at $f_N = 2$ Mc/s. This was reduced to 3·7 km by beginning the analyses from a starting height $h_s$ calculated as described above. In 230 Watheroo records which were analysed, a mean difference of 17 km was reduced to 4 km. The simple procedure suggested will thus reduce the errors, caused by the presence of ionization beneath the night-time $F$-layer, by about 5 times at frequencies above 1 Mc/s.

The difference between the solid and dotted curves in Fig. 2 gives the correction to the results of the normal ordinary ray analysis. The magnitude of the correction at any frequency is proportional to the value of $h_x' - h'_{\min}$. The values given in Table 1 can therefore be used to determine the size of this correction under any required conditions, and for any value of the magnetic dip angle $I$.

Table 1. Real height corrections in km for $h_x'(f_x) - h_0'(f_0) = 1$ km, where $f_x = 2f_H, f_0 = 1{\cdot}4f_H$

| $I$ | $f_N/f_H$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0·5 | 0·6 | 0·8 | 1·0 | 1·5 | 2 | 3 | 4 | 6 | 10 |
| 0° | 5·43 | 3·95 | 2·67 | 2·05 | 1·32 | 1·00 | 0·67 | 0·50 | 0·34 | 0·25 |
| 20° | 5·16 | 3·81 | 2·60 | 2·01 | 1·30 | 0·98 | 0·65 | 0·48 | 0·32 | 0·25 |
| 40° | 4·88 | 3·74 | 2·60 | 2·01 | 1·30 | 0·95 | 0·61 | 0·46 | 0·31 | 0·24 |
| 60° | 4·65 | 3·72 | 2·63 | 2·03 | 1·30 | 0·93 | 0·59 | 0·44 | 0·29 | 0·23 |
| 80° | 4·48 | 3·72 | 2·70 | 2·07 | 1·31 | 0·91 | 0·57 | 0·43 | 0·28 | 0·22 |

The frequency $f_0$ at which the virtual height of the ordinary ray is measured need not be the same as the lowest observed frequency $f_{\min}$. Calculations with other values of $f_0$ and $f_x$ [such that $f_0^2 = f_x(f_x - f_H)$] show that, for a given value of $h_x'(f_x) - h_0'(f_0)$, the corrections are approximately proportional to $(f_x - f_H)^2$. For $f_0 = 1{\cdot}4f_H$, the initial correction is equal to $(5{\cdot}48 - \sin I)$ within 1 per cent (Table 1). So it is found that the starting correction required at 0·5 Mc/s is given generally by

$$h'_{\min} - h_s = 0{\cdot}74(5{\cdot}48 - \sin I)(f_x - f_H)^2[h_x'(f_x) - h_0'(f_0)].$$

If the real height calculations are started from 1 Mc/s (instead of 0·5 Mc/s) the starting correction required at this frequency is:

$$h'_{\min} - h_s = 0{\cdot}36(5{\cdot}48 - \sin I)(f_x - f_H)^2[h_x'(f_x) - h_0'(f_0)].$$

These relations are accurate to within a few per cent for all practical values of $f_H$, and for values of $f_x$ less than 3 Mc/s. Larger values of $f_x$ cannot be used in practice, since the values of $h_x'(f_x) - h'_0(f_0)$ become too small to measure accurately. In the routine analysis of ionospheric records the value of $h_x'$ is measured at the lowest frequency $f_x$ at which the extraordinary ray trace is visible. The corresponding value of $h_0'$ is then found by interpolating in the measured ordinary ray heights, and the initial real height $h_s$ is calculated from the above relations.

When the extraordinary ray trace cannot be distinguished at frequencies below about 3 Mc/s, an estimated value of $h_s$ should be used. This is obtained from other records taken at the same hour on different days. Its use gives reasonably accurate results since it is found that the values of $h'_{\min}$ are largest when the amount of underlying ionization is greatest, so that the values of $h_s$ do not vary greatly from day to day.

*Acknowledgements*—This work was carried out at the Cavendish Laboratory, during the tenure of an Overseas Scholarship awarded by the Commissioners for the Exhibition of

1851. Thanks are due to the Director of the Cambridge University Mathematical Laboratory for permission to use the electronic computor EDSAC II in these calculations.

*Seagrove Radio Research Station*
*University of Auckland, New Zealand*
J. TITHERIDGE

REFERENCES

TITHERIDGE J. 1959a *J. Atmosph. Terr. Phys.* **17**, 110.
TITHERIDGE J. 1959b *J. Atmosph. Terr. Phys.* **17**, 126.

---

## The simultaneity of geomagnetic sudden impulses

(*Received* 20 *November* 1960)

INVESTIGATIONS on the degree of simultaneity of the geomagnetic-storm sudden commencement have recently been made by GERALD (1959) and WILLIAMS (1960) on the basis of IGY data, and their results were somewhat different. The purpose of this note is to find a key for the solution of the problem of whether or not geomagnetic sudden impulses and sudden commencements are essentially the same, and if they are the same, how they occur.

Table 1. Location of stations and the start time of sudden impulses

| | Geographic latitude | Geographic longitude | Oct. 1957 $14^d\ 13^h$ | Feb. 1958 $12^d\ 05^h$ | Jul. 1958 $21^d\ 19^h$ | Nov. 1958 $11^d\ 01^h$ |
|---|---|---|---|---|---|---|
| Leirvogur | N 64° 11′ | W 21° 42′ | $18^m\ 31^s$ | $55^m\ 06^s$ | $25^m\ 57^s$ | $28^m\ 41^s$ |
| Nurmijärvi | N 60° 30′ | E 24° 39′ | | | | $28^m\ 48^s$ |
| Lovö | N 59° 21′ | E 17° 50′ | $19^m\ 36^s$ | $55^m\ 15^s$ | | |
| College | N 64° 18′ | W 147° 50′ | | $55^m\ 40^s$ | $26^m\ 09^s$ | $28^m\ 33^s$ |
| Big Delta | N 64° 09′ | W 145° 51′ | | $55^m\ 04^s$ | $26^m\ 11^s$ | |
| Point Barrow | N 71° 20′ | W 156° 46′ | $19^m\ 44^s$ | $55^m\ 04^s$ | $26^m\ 12^s$ | $28^m\ 29^s$ |
| Healy | N 63° 00′ | W 151° 00′ | | $55^m\ 05^s$ | $26^m\ 14^s$ | $28^m\ 46^s$ |
| Sitka | N 57° 03′ | W 135° 20′ | $19^m\ 22^s$ | $55^m\ 16^s$ | $26^m\ 15^s$ | $29^m\ 00^s$ |
| Fredericksburg | N 38° 12′ | W 77° 22′ | $19^m\ 42^s$ | $55^m\ 19^s$ | $26^m\ 10^s$ | $29^m\ 24^s$ |
| Tucson | N 32° 14′ | W 110° 57′ | $19^m\ 52^s$ | $55^m\ 20^s$ | $26^m\ 20^s$ | $29^m\ 30^s$ |
| Shimosato | N 33° 35′ | E 135° 56′ | | $55^m\ 27^s$ | $26^m\ 21^s$ | $29^m\ 19^s$ |
| Honolulu | N 21° 18′ | W 158° 06′ | $20^m\ 06^s$ | $55^m\ 31^s$ | $26^m\ 24^s$ | $29^m\ 24^s$ |
| Koror | N 07° 16′ | E 134° 32′ | $20^m\ 13^s$ | $55^m\ 26^s$ | $26^m\ 17^s$ | $29^m\ 28^s$ |
| Guam | N 13° 27′ | E 144° 45′ | $20^m\ 14^s$ | $55^m\ 26^s$ | | $29^m\ 29^s$ |
| Wilkes | N 66° 24′ | W 110° 24′ | $19^m\ 58^s$ | $55^m\ 13^s$ | $25^m\ 52^s$ | $29^m\ 02^s$ |
| Little America | S 77° 00′ | W 164° 00′ | | $55^m\ 00^s$ | $25^m\ 52^s$ | $28^m\ 30^s$ |
| Apia | S 13° 48′ | W 171° 49′ | | $55^m\ 29^s$ | | |
| Scott Base | S 77° 51′ | W 166° 47′ | | | | $28^m\ 57^s$ |

We used copies of quick-run magnetograms during the IGY from various stations for four sudden impulses which occurred on 14 October 1957, 12 February, 21 July and 11 November 1958, because they could be easily measured. The stations are: Leirvogur, Nurmijärvi, Lovö, College, Big Delta, Point Barrow, Healy, Sitka, Fredericksburg, Tucson, Shimosato, Honolulu, Koror, Guam, Wilkes, Little America, Apia and Scott Base. Measurements were confined to $H$-(or $X$-) traces alone, and when $H$-traces were unreliable, $D$-traces were used. In graphs (a)–(d) of Fig. 1 the order of occurrence of sudden impulses is plotted



VOL. 20 1961

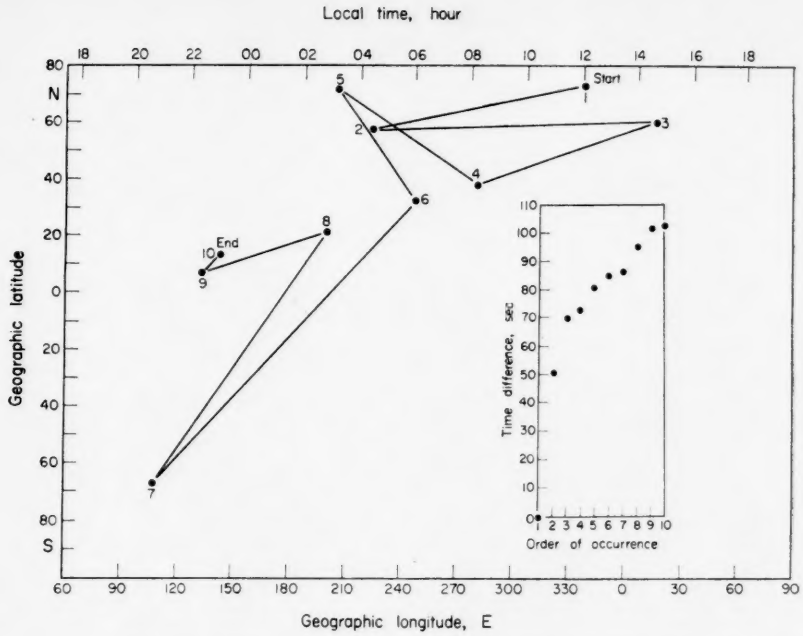


Fig. 1(a). Order of occurrence for geomagnetic sudden impulses on 14 October 1957, 1320 GMT.

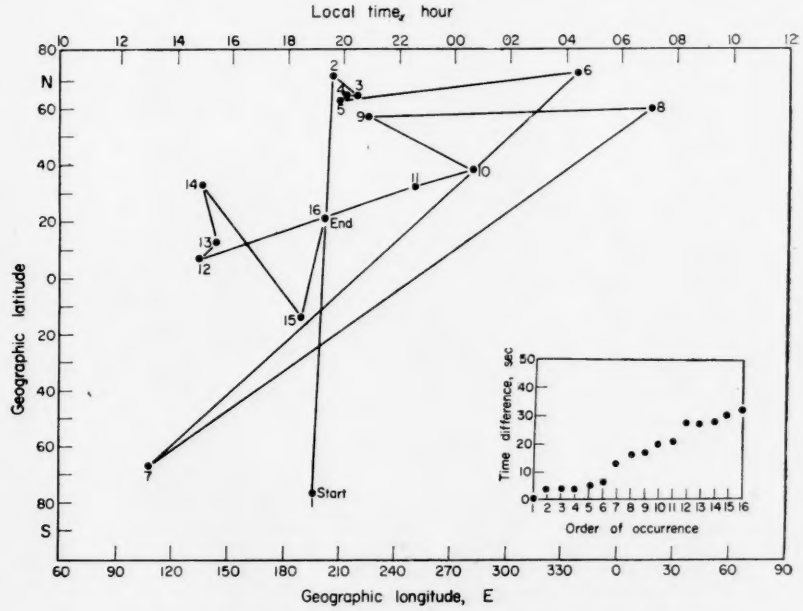


Fig. 1(b). Order of occurrence for geomagnetic sudden impulses on 12 February 1958, 0555 GMT.

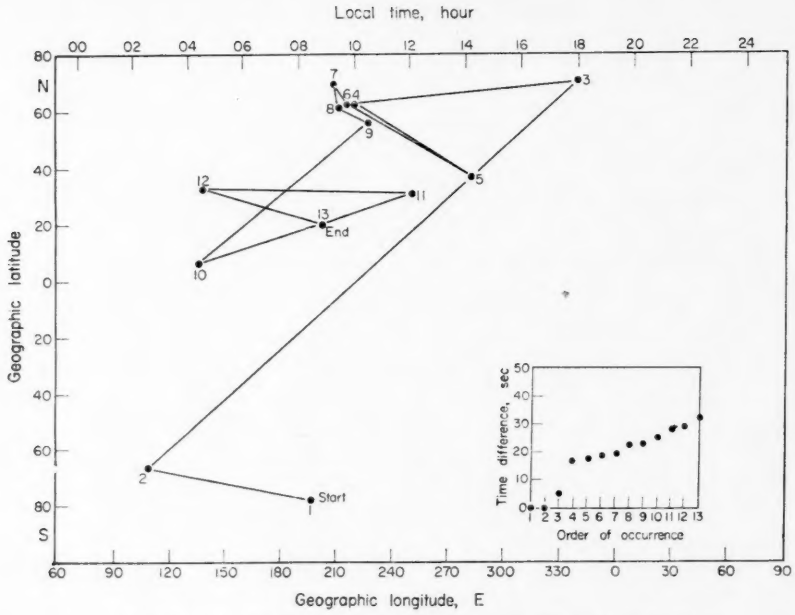


Fig. 1(c). Order of occurrence for geomagnetic sudden impulses on 21 July 1958, 1926 GMT.

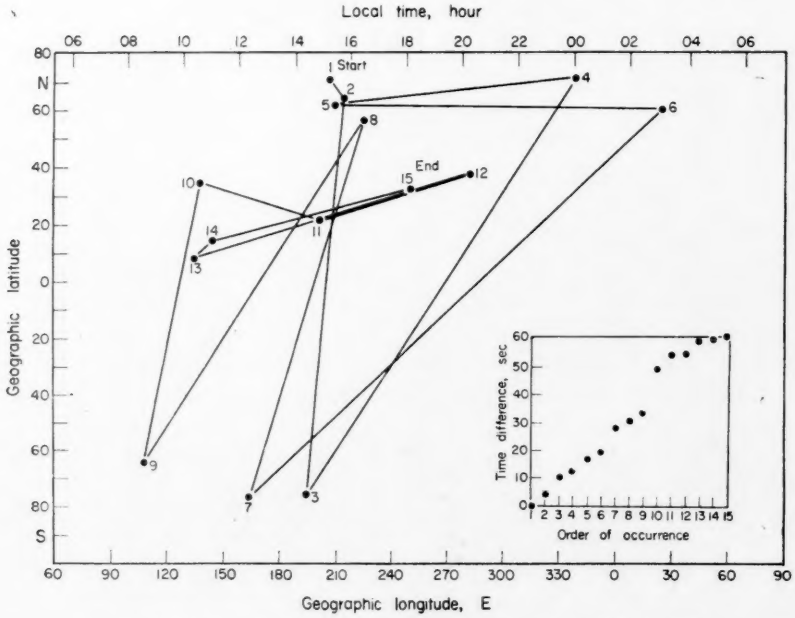


Fig. 1(d). Order of occurrence for geomagnetic sudden impulses on 11 November 1958, 0129 GMT.



VOL. 20 1961

in a form similar, for the convenience of comparison, to that illustrated by WILLIAMS (1960). Geographic latitude, geographic longitude and the local time are given around each graph. On the right of the graphs are plots of the order of occurrence against time.

These analyses of the data yielded the following results:

(1) Time differences of sudden impulses around the earth were within 1 min;
(2) The sudden impulses always occurred first in high latitudes;
(3) The average propagation velocity of sudden impulses between Honolulu and Koror was about 1300 km/sec.

These results are very similar to those obtained by WILLIAMS (1960) for sudden commencements of magnetic storms. Further study is in progress, and its result will be published later.

*Acknowledgement*—The authors wish to thank Miss F. KAWAI for her assistance.

*Geophysical Institute*
*Kyoto University*
*Kyoto, Japan*

M. YAMAMOTO
H. MAEDA

REFERENCES

GERARD V. B. 1959 *J. Geophys. Res.* **64**, 593.
WILLIAMS V. L. 1960 *J. Geophys. Res.* **65**, 85.

---

## On the large scale regions of irregularities producing scintillation of signals transmitted from earth satellites

(*Received 5 December* 1960)

RECORDINGS of the signal strength of radio transmissions from earth satellites show two kinds of fading. One is a very regular slow fading with neat maxima and minima caused by Faraday-rotation and the rotation of the satellite. The other is much faster, irregular and shows no periodicity. This fading is caused by irregularities in the ionosphere and is often called scintillation.

KENT (1959) recorded the strength of the 40 Mc/s signal from the first Russian satellite at Cambridge. For the period of observation he never observed scintillation of the signal when the satellite was to the south of Cambridge. He concluded that the irregularities producing the scintillation had not occurred south of about 50° N.

MAWDSLEY (1960) proposed that the reason no scintillation was observed south of Cambridge is connected with the scattering properties of the irregularities, these being elongated along the geomagnetic lines of force, and having a minimum scattering efficiency for waves travelling in their direction of elongation.

At Kjeller (60°N, 11°E) signal strength recordings have been made on transmissions on 20 Mc/s from 1958 $\delta$ II, and on 108 Mc/s from 1960 $\eta$ II. An examination of about fifty recordings of transits of 1958 $\delta$ II has given the results shown in the figure. The criterion used for distinguishing scintillation is the absence of regular low speed fading with neat maxima and minima. It is seen that there is no marked difference in the frequency of occurrence of scintillation when the satellite was to the north or to the south of Kjeller.

It may be noted that normally the signal showed either Faraday rotation or scintillation for the whole observation period of the transit.

It may also be pointed out that at the time these recordings were made the satellite was moving at a height of about 1000 km while the irregularities have been found to lie at a height of 300–500 km. Thus the irregularities producing the scintillation are situated about halfway between satellite and observing point.

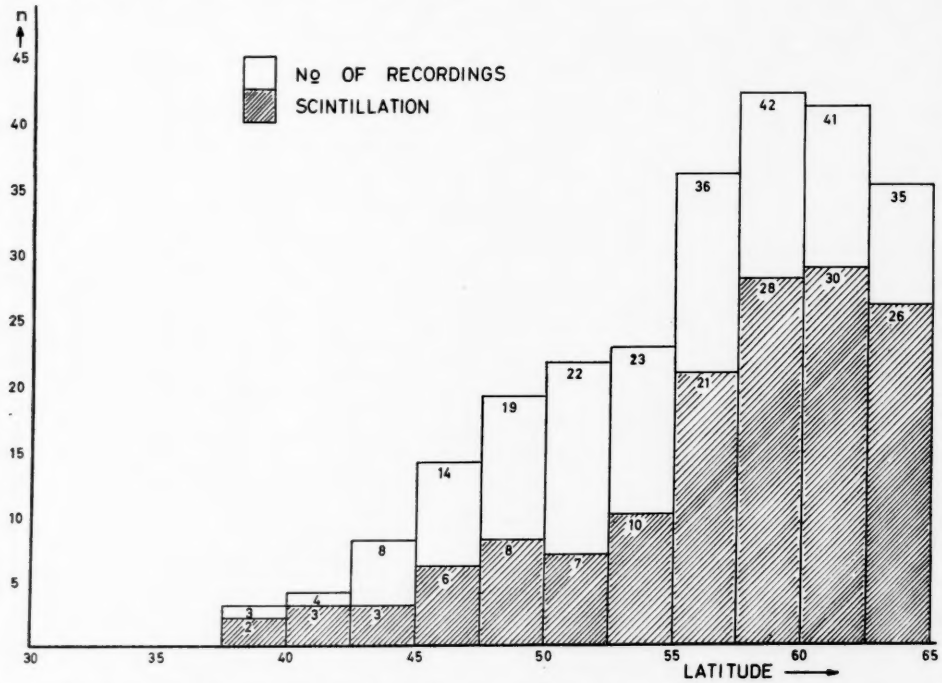


Fig. 1. Number of times scintillations of the satellite were observed compared with the total number of observations for different latitudes of the transmitting satellite.

Recordings of satellite transmissions have also been made at Tromsö (70°N, 19°E). Here one nearly always observes scintillation of the signal both on 20 and 108 Mc/s. However, none of the satellites observed have passed to the north of Tromsö.

The results here presented indicate that the effect proposed by MAWDSLEY is not strong enough to blot out the scintillation of satellites to the south of Tromsö or Kjeller.

It is concluded that the aspect sensitivity of the scattering is a minor effect in the control of scintillation.

*Acknowledgements*—The research reported in this note has been sponsored by the Electronics Research Directorate of Cambridge Research Center, Air Research and Development Command, United States Air Force, through the European Office, ARDC, under contract AF 61(052)–08.

J. FRIHAGEN
J. TRÖIM

*Norwegian Defense Research Establishment*
*Division for Telecommunication*
*Kjeller, Lillestrom, Norway*

REFERENCES

KENT G. S. 1959 *J. Atmosph. Terr. Phys.* **16,** 10.
MAWDSLEY J. 1960 *J. Atmosph. Terr. Phys.* **18,** 344.



VOL. 20 1961

# On the excitation of the First Positive $N_2$ bands in aurorae*

(*Received* 8 *December* 1960)

As pointed out by BATES (1960) the excitation of the First Positive $N_2$ bands in aurorae may well be explained by the process

$$N_2(X^1\Sigma) + e \rightarrow N_2(B^3\Pi) + e. \quad (1)$$

Other processes, such as

$$N_2^+ + O^- \text{ (or } O_2^-) \rightarrow N_2(B^3\Pi) + O \text{ (or } O_2) \quad (2a,b)$$

have also been thought to be operative, and were recently discussed by MALVILLE (1959), who thinks that this latter process may be particularly effective in type B aurora (red lower border due to enhancement of the First Positive bands). Processes (2a,b) should give a time delay, not less than 1 sec, between the ionization process and the emission of quanta in the First Positive system. It should be possible to observe the corresponding time delay between the emission of the First Negative $N_2^+$ band and the First Positive bands when the aurora varies rapidly in intensity. With process (1) the time delay should be very small, since the $N_2^+$ bands almost certainly are excited simultaneously with the ionization, the two excitation processes being almost simultaneous (cf. OMHOLT, 1959a).

In order to clarify this problem the intensities of the First Positive $N_2$-bands and the First Negative $N_2^+$ bands emitted simultaneously from aurorae were recorded with a photo-electric two-colour filter photometer. The photometer is the same as used previously in a similar study of the [OI] $\lambda$ 5577 line and the First Negative bands (OMHOLT, 1959a). The (0, 1) Negative band ($\lambda$ = 4278 Å) and the (5, 2) Positive band ($\lambda$ = 6710 Å) were isolated with multilayer interference filters (centred at 4278 and 6690 Å and with half-widths approximately 50 Å), and recorded by EMI 9558 photomultipliers. The photometer covers a circular field in the sky, the diameter of the field being about 0·8° only. The intensity fluctuations of the two emissions were finally recorded with an Evershed pen recorder and, when rapid fluctuations occured, also with an oscilloscope fitted with a Shackman 35 mm film camera.

Records were obtained of all kinds of variable aurorae, including type B. The records show clearly that the two measured bands fluctuate in phase, and that there is no detectable, systematic time delay between the two emissions, even in rapidly fluctuating and pulsating aurorae. Some phase differences occur, but these are very small and most likely due to random variations in the photo current. To give a safe upper limit we may say that the average delay between the emission of the two bands is less than 0·1 sec. Since the processes (2a, b) definitely must give an appreciable delay, we may conclude that process (1) is probably the dominating one in exciting the First Positive bands, as the similar process (simultaneous ionization and excitation) is in exciting the First Negative bands. It should be pointed out, however, that only one record of type B aurora has been obtained so far, so that the result is not conclusive for this type of aurora.

MALVILLE (1959) discusses processes (2a, b) and concludes that the time constant of this excitation mechanism, including the preceding formation of negative ions (and consequently the discussed time delay) may be as low as 1 sec. However, he has probably overestimated the electron production rate in the aurora by a factor of 100, because the emission rate of the First Negative band $\lambda$ 3914 in an aurora of brightness III is ten times lower than quoted by him [Seaton's u.v. factor is only for correcting previous, wrong intensities, cf.

*The work reported in this note was supported in part by the Geophysics Directorate of the U.S. Air Force, ARDC, under contract AF61(514)–1123 through the European Office, Brussels.

also BATES (1960) and OMHOLT (1959b)], and it is hardly realistic to take the length of the emitting column to be less than 10 km. With these corrections we obtain an electron density of the order $10^6$ rather than $10^7$ cm$^{-3}$. This is also in agreement with direct measurements of the relation between the auroral luminosity in zenith and the critical frequency of the $E_s$-layer accompanying the aurora (cf. OMHOLT, 1959b). Radio echoes from aurorae on about 100 Mc/s are most certainly a backscatter phenomena, and yield little or no information about the absolute electron density in aurorae. Following MALVILLE's arguments, the time delay between the First Negative and the First Positive bands then should rather be 10 sec or more if the processes (2a, b) are dominating. It should be possible to observe such a delay visually, and it has also been looked for without detecting any effect.

One main point in MALVILLE's discussion is that $\lambda$, the negative ion to electron ratio, is high in type B aurora. The measurements of the correlation between the auroral luminosity and critical frequencies in the $E_s$-layer yield an apparent recombination coefficient of about $10^{-6}$ cm$^3$ sec$^{-1}$, a value which is difficult to explain if one does not assume that $\lambda$ is high also in ordinary aurorae (cf. OMHOLT, 1959b). If this is so, the processes (2a, b) should be operative in ordinary aurorae also, if they are in type B aurorae, since then negative ions are abundant in all types. This is again an argument against these processes, although one should bear in mind that the kind of negative ion dominating may vary from one aurora to another, since the height interval where aurorae occur includes the $O_2$–O transition region ($O_2^-$–$O^-$). If the effective rate coefficients for the two processes (2a) and (2b) are very different, the importance of excitation through negative ions may be different above and below the $O_2$–O transition region (corresponding to ordinary and type B aurora).

If excitation through the processes (2a, b) is a slow process, giving a rather long delay between the primary process and the excitation of the First Positive bands, one would expect that the light excited in this way would not occur in the rapidly varying aurora, but as a relatively weak afterglow of long duration and with large time constant. Then, what we measure in the rapidly varying aurora would be only that fraction of the light which is excited through process (1). But there appears to be no systematic difference in the intensity ratio between the First Positive bands and the First Negative bands emitted from varying and quiet aurorae, and no sign of such an afterglow have been positively found, although it cannot be excluded in a few cases. It has neither been observed visually nor been reported by other observers. The ratio of the measured intensities does vary from one aurora to another, with as much as a factor of 2, but this variation is apparently not systematic, and is probably due to variations in the energy spectrum of the active electrons (cf. BATES, 1960). The absolute value of this intensity ratio has not been measured, but the scatter in this ratio is greater among weak aurorae. Also, the intensity ratio is fairly constant in one particular auroral feature, although the absolute intensities may vary considerably.

In conclusion we may say that excitation through mutual neutralization of negative ions and $N_2^+$ ions probably contributes to less than 10 per cent of the total excitation of the First Positive bands in almost all types of aurora. This seems also to be the case for type B aurora, but the conclusion is much less certain in this case. A more detailed account of the work, which will be continued, will be given elsewhere later.

A. OMHOLT

*The Norwegian Institute of Cosmic Physics, Blindern*
*and The Institute of Theoretical Astrophysics*
*University of Oslo, Blindern, Norway*

## REFERENCES

BATES D. R. 1960 *Physics of the Upper Atmosphere* (Ed. by J. A. RATCLIFFE) Chap. 7. Academic Press, New York.



VOL. 20 1961

| | | |
|---|---|---|
| MALVILLE J. M. | 1959 | *J. Atmosph. Terr. Phys.* **16,** 59. |
| OMHOLT A. | 1959a | *Geofys. Publ.* **21,** No. 1. |
| OMHOLT A. | 1959b | *Electromagnetic Wave Propagation* (Ed. by M. DESIRANT and J. L. MICHIELS) p. 971. Academic Press, New York. |


OL.
20
961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 220 to 223. Pergamon Press Ltd. Printed in Northern Ireland

# BOOK REVIEWS

**Physics and Medicine of the Atmosphere and Space.** Edited by O. O. BENSON, Jr. and H. STRUGHOLD. John Wiley, New York, London, 1960. xviii + 645 pp., £5.

DURING November 1959 there was held in San Antonio the Second International Symposium on Physics and Medicine of the Atmosphere and Space, under the sponsorship of the U.S. Air Force School of Aviation Medicine with the assistance of the Southwest Research Institute. The present volume contains over forty papers presented at that symposium. Each paper is by a different authority, on either the physical or physiological aspects of space travel, and yet, as a result of careful editing, the whole volume is found to possess a cohesive design. The reader whose interests are in physical science cannot fail to be impressed, and may well be surprised, by the enormous amount of progress which has been made in the study of the human environmental factors which man will encounter during space journeys. Such progress is described here in most readable fashion.

Of the physical articles likely to be of most pertinent interest to readers of this Journal I might list the following: "On the Radiation Hazards of Space Flight", by JAMES A. VAN ALLEN; "Aeronomic Chemical Reactions", by M. NICOLET; "Meteoric Material in Space", by FRED L. WHIPPLE; "Effects of Interplanetary Dust and Radiation Environment on Space Vehicles" by S. FRED SINGER; "The Electromagnetic Environment of the Atmosphere and Nearer Space", by WALTER DIEMINGER; "Upper Atmospheric Properties based on Rocket and Satellite Data", by H. KORF KALLMANN; "Composition of the Upper Atmosphere", by JOHN W. TOWNSEND; and "The Physics of the Sun", by WALTER ORR ROBERTS. The names of these authors constitute in themselves a guarantee of the authoritative nature of their contributions to this important source book of data on space travel.

E. V. APPLETON

---

G. E. DUNN and B. I. MILLER: **Atlantic Hurricanes.** Louisiana State University Press, 1960. 326 pp., $10.

FOR A POPULAR BOOK, addressed to the layman, this volume is remarkable, both for the thoroughness with which every characteristic and effect of Atlantic hurricanes is examined and for the clarity of the accounts it gives of the physical processes involved in the formation, development and dissipation of these catastrophic storms (17,000 lives have been lost in hurricanes in America since 1900) so far as they are known. It is thoroughly up to date, containing accounts of quite recent work and a photograph of a typhoon in Australian waters on 10 April 1960, provided by the satellite Tiros I. Side by side at the end of the book are the unusual combination of a glossary of meteorological terms for the uninitiated reader and a list of references to over 100 scientific papers. Each will be used; for this is a book that will be read with interest by the layman, especially if he lives within the orbit of its subject, yet it will also be consulted by the specialist, for its content of statistical data qualifies it to serve also as a work of reference.

J. PATON



VOL. 20 1961

W. YOURGRAU and S. MANDELSTAM: **Variational Principles in Dynamics and Quantum Theory** (Second Edition). Pitman, London, 1960. xi + 180 pp., 32s.6d.

THE FIRST EDITION of this stimulating book was reviewed in this journal by L. ROSENFELD in 1955 (*J. Atmosph. Terr. Phys.* **7,** 110). The new edition differs from the earlier one by the inclusion of some significant new material: two chapters, which deal with the formulations of quantum theory associated mainly with the names of BOHR and SOMMERFELD and of SCHRÖDINGER respectively, now have postscripts suggested by remarks addressed to the authors by Professor SCHRÖDINGER after the publication of the previous edition; there is a new fifteen-page chapter on "The principles of FEYNMAN and SCHWINGER in quantum mechanics"; and a paper by YOURGRAU and RAW on "Variational principles and chemical reactions", first printed in *Nuovo Cimento* in 1957, appears as an appendix.

Although the laws of classical mechanics and electromagnetism can be expressed through variational principles, and although the Hamiltonian formulation of these laws is the basis of the usual approach to the more general laws of quantum mechanics, variational principles do not have the prominent place in quantum theory that they occupy in much of classical physics. Within the past 12 years, however, FEYNMAN and SCHWINGER have shown, by methods which are superficially very different but are in fact equivalent, that the laws of quantum mechanics can be formulated with a new succinctness in terms of an action principle which reduces, in the classical limit, to the variational principle of least action. The authors describe and discuss this important development in quantum theory, in the general setting of their concise review of the place and history of variational principles in physics. This is a valuable addition to a book which has already been welcomed for the cold clear light which it shines on topics which have too often in the past been viewed through a haze of mysticism.

R. M. SILLITTO

---

C. H. B. PRIESTLEY: **Turbulent Transfer in the Lower Atmosphere.** The University of Chicago Press and Cambridge University Press, 1959. vii + 130 pp., 28s.

THE STUDY of atmospheric turbulence is among the most highly specialized and, by repute, among the most "difficult" branches of meteorology. Such study tends at present to be restricted to small-scale phenomena: apart from the intrinsic importance and direct applications of results obtained at this scale, however, there is general agreement that such studies may bear fruit in the solution of problems relating to much larger-scale phenomena, since the transfer of atmospheric properties implicit in the "general circulation" of the atmosphere is mainly effected by the large-scale eddies known as depressions and anticyclones.

The book under review is concerned with an explanation of the detailed mechanism of the vertical transfer of momentum, heat and matter which is effected in fine-scale atmospheric turbulence. It naturally leans heavily on the measurement techniques and results obtained by the team of Australian workers which the author heads, but is by no means neglectful of the experimental and theoretical work which has been done elsewhere. The scope of the book is best illustrated by the chapter titles which, following a brief introduction, are: "The Eddy Flux and its Measurement"; "The Shearing Stress and the Wind Profile"; "Heat Convection and the Temperature Profile"; "The Spectrum of Turbulence and the Structure of Free Convection"; "Theories of Buoyant Motion"; "Evaporation"; and "Evolutionary Aspects of Energy Transfer".

The book was written, not as a text-book, but as an extension of a series of lectures delivered by the author on a visit to Chicago. It is an obvious "must" for workers in the specialized fields indicated in the chapter titles. It represents, however, a piece of work so remarkable in its conciseness of writing and in its treatment of the various topics as to be unsuitable for any non-specialist who does not first make himself familiar with the contents of O. G. SUTTON's *Micrometeorology* which is one of the few textbooks in any language in this field of investigation.

D. H. MCINTOSH

**Space Research**—*Proceedings of the First International Space Symposium*. Edited by Dr. H. KALLMAN BIJL. Interscience, New York, 1960. 1195 pp., $24.00.

THIS VOLUME contains about one hundred scientific articles, printed on well over a thousand pages; it is also over five pounds in weight! It is well produced typographically; but its high price may well curtail its sale to individuals. The library of every scientific institution in the world should, however, secure a copy for its readers.

At the end of the International Geophysical Year 1957/58, the International Council of Scientific Unions, which had itself sponsored the I. G. Y., established a new Committee on Space Research (COSPAR), under the distinguished Presidency of Professor H. C. VAN DE HULST and Vice-Presidencies of Professor A. A. BLAGONRAVOV and Dr. R. PORTER. By the terms of its charter, COSPAR was enjoined:

> "to further on an international scale the progress of all kinds of scientific investigations which are carried out with the use of rockets or rocket-propelled vehicles. COSPAR shall be concerned with fundamental research. It will not normally concern itself with such technological problems as propulsion, construction of rockets, guidance and control."

Among its many other activities COSPAR organized the first "International Space Science Symposium", which was held at Nice during January 1960. The subjects discussed at the Symposium were limited to scientific research and were, for the most part, of a non-biological nature. As a result, the formulation of plans for human space travel was not before the Conference. Instead, attention was concentrated on observations already achieved by way of instrumented earth satellites, space probes and rockets during the two years since the first sputnik was launched on 4 October 1957. Papers were presented by space scientists from both the U.S.S.R. and the U.S.A.; but there were also contributions from workers in other countries who, most enterprisingly, had made observations on space vehicles traversing their own skies.

The present volume contains the scientific papers presented at the Nice Conference, without a record of the discussion. The contributions are all in English, though, in every case, there is also given a translation of the abstract in Russian. In some cases a French abstract is also provided. The papers are grouped under seven headings as follows: (1) The Earth's Atmosphere (18 papers); (2) The Ionosphere (13 papers); (3) Tracking and Telemetering (15 papers); (4) Solar Radiation (16 papers); (5) Cosmic Radiation (22 papers); (6) Interplanetary Dust (7 papers); and (7) The Moon and the Planets (8 papers). It should be stressed that all the contributions are severely scientific, there being no attempt to interpret the progress of space science to the layman.

It is obviously impossible to traverse here the subject-matter of these hundred papers. But it should be said that the publication of this volume is a landmark in the documentation of space science. The triumph of the organization of the Nice Conference must obviously be repeated, at appropriate intervals, by COSPAR. If the services of the same editor (Dr. H. KALLMAN BIJL) can be secured for further Conference publications, so much the better.

E. V. APPLETON

---

**Electromagnetic Wave Propagation.** Edited by M. DESIRANT and J. L. MICHIELS. Academic Press, London, 1960. xii + 730 pp., 157s.

IN planning the 1958 International Exhibition in Brussels the organizers sought to illustrate how the progress of science has contributed to the welfare of mankind. There was, therefore, associated with the Exhibition a number of international scientific conferences, to which representatives of Universities, Government Departments and Industry were invited. The fifty-four papers presented at one such conference, on radio wave propagation, constitute the material of this volume. In order to facilitate early publication, the editors explain, no attempt



VOL. 20 1961

has been made to standardize notations and nomenclatures. For the same reason, presumably, a few of the papers have been printed in their original language, French or German; though by far the greatest number are in English. It should also be added that there is no page index of either names or subjects.

Speaking generally, it can be said that the papers in this volume constitute a contribution to radio technology rather than to radio science, so that their appeal will be to the communication engineer rather than to the radio-physicist. They deal with the ways radio waves travel, usefully, over long and short distances; and touch on atmospheric physics only when concerned with those properties of the terrestrial atmosphere which facilitate, or discourage, such travel.

E. V. Appleton

OL.
20
961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 224 to 227. Pergamon Press Ltd. Printed in Northern Ireland

# Indices of geomagnetic activity

of the Observatories HARTLAND (*Ha*), ESKDALEMUIR (*Es*) and LERWICK (*Le*)

September–December 1960

The figures given on pages 224 to 227 represent the $K$-indices for three-hour intervals, beginning with 00–03 hrs for the first and ending with 21–24 hrs for the eighth figure.

*September* 1960

| *Day* | *Ha* Range for $K = 9 : 500\gamma$ | | | *Es* Range for $K = 9 : 750\gamma$ | | | *Le* Range for $K = 9 : 1000\gamma$ | | | *Day* |
|---|---|---|---|---|---|---|---|---|---|---|
| | *K-Indices* | | *Sum* | *K-Indices* | | *Sum* | *K-Indices* | | *Sum* | |
| 1 | 0000 | 1101 | 3 | 1122 | 1211 | 11 | 1112 | 1102 | 9 | 1 |
| 2 | 0111 | 3345 | 18 | 1222 | 3345 | 22 | 1122 | 3336 | 21 | 2 |
| 3 | 5452 | 3321 | 25 | 4554 | 3332 | 29 | 7453 | 3332 | 30 | 3 |
| 4 | 3454 | 5557 | 38 | 4455 | 6658 | 43 | 2455 | 6679 | 44 | 4 |
| 5 | 5666 | 3555 | 41 | 7666 | 4555 | 44 | 8766 | 4557 | 48 | 5 |
| 6 | 4332 | 1134 | 21 | 5344 | 1244 | 27 | 6452 | 2236 | 30 | 6 |
| 7 | 3223 | 5343 | 25 | 3444 | 5344 | 31 | 4223 | 4343 | 25 | 7 |
| 8 | 4242 | 2330 | 19 | 4344 | 4432 | 28 | 4333 | 3431 | 24 | 8 |
| 9 | 3322 | 2223 | 19 | 4332 | 2233 | 22 | 4422 | 2233 | 22 | 9 |
| 10 | 2213 | 2332 | 18 | 2213 | 3333 | 20 | 3112 | 3433 | 20 | 10 |
| 11 | 4321 | 2333 | 21 | 4322 | 3333 | 23 | 4421 | 2334 | 23 | 11 |
| 12 | 3221 | 1321 | 15 | 3232 | 2322 | 19 | 3222 | 3221 | 17 | 12 |
| 13 | 4432 | 2134 | 23 | 4433 | 3234 | 26 | 4542 | 2225 | 26 | 13 |
| 14 | 5322 | 1200 | 15 | 4322 | 2211 | 17 | 6332 | 1211 | 19 | 14 |
| 15 | 0111 | 1111 | 7 | 1222 | 2222 | 15 | 1121 | 2111 | 10 | 15 |
| 16 | 1001 | 2110 | 6 | 2111 | 2211 | 11 | 2011 | 2111 | 9 | 16 |
| 17 | 0101 | 2243 | 13 | 1211 | 2344 | 18 | 1111 | 2244 | 16 | 17 |
| 18 | 4323 | 2112 | 18 | 4323 | 3212 | 20 | 4322 | 2112 | 17 | 18 |
| 19 | 1011 | 2000 | 5 | 2122 | 2210 | 12 | 2111 | 1100 | 7 | 19 |
| 20 | 0010 | 1123 | 8 | 0111 | 1223 | 11 | 1011 | 1114 | 10 | 20 |
| 21 | 3012 | 2233 | 16 | 2012 | 2233 | 15 | 4112 | 2234 | 19 | 21 |
| 22 | 3110 | 2233 | 15 | 3221 | 3333 | 20 | 3111 | 2233 | 16 | 22 |
| 23 | 2333 | 1124 | 19 | 3333 | 3223 | 22 | 3333 | 1215 | 21 | 23 |
| 24 | 4343 | 2341 | 24 | 4343 | 3332 | 25 | 6533 | 2442 | 29 | 24 |
| 25 | 1001 | 1132 | 9 | 1112 | 1132 | 12 | 0011 | 1132 | 9 | 25 |
| 26 | 3021 | 1233 | 15 | 2122 | 1233 | 16 | 2222 | 1136 | 19 | 26 |
| 27 | 4332 | 1122 | 18 | 4333 | 1232 | 21 | 6332 | 1222 | 21 | 27 |
| 28 | 0112 | 1201 | 8 | 1123 | 2211 | 13 | 0112 | 1211 | 9 | 28 |
| 29 | 1021 | 4224 | 16 | 1122 | 4324 | 19 | 1021 | 4335 | 19 | 29 |
| 30 | 4332 | 3354 | 27 | 4332 | 3334 | 25 | 5432 | 3355 | 30 | 30 |



VOL. 20 1961

*October* 1960

| *Day* | *Ha* Range for $K = 9 : 500\gamma$ | | | *Es* Range for $K = 9 : 750\gamma$ | | | *Le* Range for $K = 9 : 1000\gamma$ | | | *Day* |
|---|---|---|---|---|---|---|---|---|---|---|
| | *K-Indices* | | *Sum* | *K-Indices* | | *Sum* | *K-Indices* | | *Sum* | |
| 1 | 5544 | 4346 | 35 | 4544 | 4345 | 33 | 4543 | 4455 | 34 | 1 |
| 2 | 4443 | 4553 | 32 | 3443 | 5543 | 31 | 5444 | 5543 | 34 | 2 |
| 3 | 2224 | 3312 | 19 | 2223 | 3322 | 19 | 4223 | 2213 | 19 | 3 |
| 4 | 2012 | 3565 | 24 | 2122 | 3565 | 26 | 3012 | 3677 | 29 | 4 |
| 5 | 4432 | 3435 | 28 | 4432 | 3435 | 28 | 4432 | 3337 | 29 | 5 |
| 6 | 5656 | 6888 | 52 | 4656 | 7889 | 53 | 7665 | 7999 | 58 | 6 |
| 7 | 8756 | 6655 | 48 | 9866 | 6656 | 52 | 9866 | 7767 | 56 | 7 |
| 8 | 5334 | 3344 | 29 | 5234 | 3344 | 28 | 7433 | 3354 | 32 | 8 |
| 9 | 3443 | 4355 | 31 | 3433 | 4345 | 29 | 5542 | 4446 | 34 | 9 |
| 10 | 4111 | 0123 | 13 | 3320 | 1123 | 15 | 4411 | 0124 | 17 | 10 |
| 11 | 4232 | 2123 | 19 | 3332 | 2223 | 20 | 4332 | 3333 | 24 | 11 |
| 12 | 0012 | 1124 | 11 | 1112 | 1223 | 13 | 2022 | 0223 | 13 | 12 |
| 13 | 0111 | 1102 | 7 | 1212 | 2011 | 10 | 2111 | 1101 | 8 | 13 |
| 14 | 1011 | 0000 | 3 | 1112 | 1001 | 7 | 1011 | 0101 | 5 | 14 |
| 15 | 1222 | 3343 | 20 | 1222 | 3344 | 21 | 1112 | 3564 | 23 | 15 |
| 16 | 2221 | 2002 | 11 | 2122 | 2202 | 13 | 3111 | 2202 | 12 | 16 |
| 17 | 2112 | 2122 | 13 | 2112 | 2122 | 13 | 2111 | 2122 | 12 | 17 |
| 18 | 3321 | 3354 | 24 | 3332 | 3355 | 27 | 4432 | 3575 | 33 | 18 |
| 19 | 2312 | 2224 | 18 | 2323 | 2224 | 20 | 2332 | 1134 | 19 | 19 |
| 20 | 0221 | 1223 | 13 | 1222 | 2323 | 17 | 1212 | 2323 | 16 | 20 |
| 21 | 2111 | 1124 | 13 | 2122 | 2024 | 15 | 3121 | 2123 | 15 | 21 |
| 22 | 0000 | 0000 | 0 | 0011 | 1110 | 5 | 1011 | 0100 | 4 | 22 |
| 23 | 0000 | 1122 | 6 | 1011 | 1122 | 9 | 1000 | 1122 | 7 | 23 |
| 24 | 0011 | 4542 | 17 | 1112 | 3542 | 19 | 0001 | 3753 | 19 | 24 |
| 25 | 2244 | 4764 | 33 | 2244 | 5875 | 37 | 2245 | 5897 | 42 | 25 |
| 26 | 5434 | 4454 | 33 | 6544 | 4443 | 34 | 7545 | 5455 | 40 | 26 |
| 27 | 4324 | 3455 | 30 | 4324 | 3455 | 30 | 5423 | 3556 | 33 | 27 |
| 28 | 4534 | 3553 | 32 | 3534 | 3543 | 30 | 5544 | 4543 | 34 | 28 |
| 29 | 4334 | 4444 | 30 | 4334 | 4444 | 30 | 5233 | 4545 | 31 | 29 |
| 30 | 4334 | 3445 | 30 | 4334 | 4444 | 30 | 3334 | 4345 | 29 | 30 |
| 31 | 4333 | 3553 | 29 | 4333 | 3553 | 29 | 5333 | 3553 | 30 | 31 |

*November* 1960

| *Day* | *Ha* *Range for K = 9 : 500γ* | | | *Es* *Range for K = 9 : 750γ* | | | *Le* *Range for K = 9 : 1000γ* | | | *Day* |
|---|---|---|---|---|---|---|---|---|---|---|
| | *K-Indices* | | *Sum* | *K-Indices* | | *Sum* | *K-Indices* | | *Sum* | |
| 1 | 2222 | 2224 | 18 | 2222 | 2224 | 18 | 3222 | 2224 | 19 | 1 |
| 2 | 3222 | 3134 | 20 | 2232 | 3143 | 20 | 2232 | 3233 | 20 | 2 |
| 3 | 3212 | 2144 | 19 | 3322 | 3234 | 22 | 4322 | 3236 | 25 | 3 |
| 4 | 5433 | 4544 | 32 | 5533 | 4544 | 33 | 6653 | 6555 | 41 | 4 |
| 5 | 3321 | 2000 | 11 | 3322 | 2101 | 14 | 5322 | 2211 | 18 | 5 |
| 6 | 0111 | 2032 | 10 | 1122 | 1132 | 13 | 0112 | 2132 | 12 | 6 |
| 7 | 3211 | 1000 | 8 | 2222 | 1101 | 11 | 3321 | 1001 | 11 | 7 |
| 8 | 0011 | 0002 | 4 | 0112 | 1112 | 9 | 0011 | 1103 | 7 | 8 |
| 9 | 2222 | 1002 | 11 | 2222 | 2102 | 13 | 2212 | 1102 | 11 | 9 |
| 10 | 0022 | 2210 | 9 | 0023 | 3211 | 12 | 0012 | 2310 | 9 | 10 |
| 11 | 2333 | 3233 | 22 | 2333 | 3233 | 22 | 1223 | 3223 | 18 | 11 |
| 12 | 3012 | 5577 | 30 | 2111 | 4578 | 29 | 3111 | 4789 | 34 | 12 |
| 13 | 8799 | 8676 | 60 | 9897 | 8775 | 60 | 8999 | 8978 | 67 | 13 |
| 14 | 5333 | 3454 | 30 | 7343 | 3453 | 32 | 9333 | 2454 | 33 | 14 |
| 15 | 4332 | 5556 | 33 | 4333 | 5557 | 35 | 4332 | 6868 | 40 | 15 |
| 16 | 8844 | 5433 | 39 | 9744 | 6433 | 40 | 9945 | 5533 | 43 | 16 |
| 17 | 3222 | 2342 | 20 | 3223 | 3442 | 23 | 3222 | 2442 | 21 | 17 |
| 18 | 3201 | 1000 | 7 | 4112 | 1110 | 11 | 3201 | 0101 | 8 | 18 |
| 19 | 0212 | 3102 | 11 | 0222 | 3112 | 13 | 0111 | 2101 | 7 | 19 |
| 20 | 2212 | 3211 | 14 | 2222 | 3212 | 16 | 1111 | 2211 | 10 | 20 |
| 21 | 1353 | 4454 | 29 | 1353 | 3565 | 31 | 1343 | 4676 | 34 | 21 |
| 22 | 5333 | 2224 | 24 | 5333 | 2324 | 25 | 5532 | 2334 | 27 | 22 |
| 23 | 2231 | 1022 | 13 | 2232 | 2122 | 16 | 2231 | 2123 | 16 | 23 |
| 24 | 4221 | 2235 | 21 | 3221 | 2235 | 20 | 3221 | 1225 | 18 | 24 |
| 25 | 4344 | 4544 | 32 | 3344 | 3444 | 29 | 3544 | 3565 | 35 | 25 |
| 26 | 3322 | 2133 | 19 | 2333 | 3133 | 21 | 3332 | 2233 | 21 | 26 |
| 27 | 3322 | 3454 | 26 | 3222 | 3444 | 24 | 3222 | 2454 | 24 | 27 |
| 28 | 5432 | 3112 | 21 | 4433 | 3222 | 23 | 5432 | 3201 | 20 | 28 |
| 29 | 3122 | 1123 | 15 | 3112 | 2223 | 16 | 3112 | 1123 | 14 | 29 |
| 30 | 2311 | 2245 | 20 | 2321 | 2245 | 21 | 2210 | 1136 | 16 | 30 |



VOL. 20 1961

*December* 1960

| *Day* | *Ha* Range for $K = 9 : 500\gamma$ | | *Sum* | *Es* Range for $K = 9 : 750\gamma$ | | *Sum* | *Le* Range for $K = 9 : 1000\gamma$ | | *Sum* | *Day* |
|---|---|---|---|---|---|---|---|---|---|---|
| | *K-Indices* | | | *K-Indices* | | | *K-Indices* | | | |
| 1 | 5643 | 4556 | 38 | 5654 | 4656 | 41 | 7854 | 5766 | 48 | 1 |
| 2 | 3333 | 4344 | 27 | 3333 | 4344 | 27 | 3333 | 4344 | 27 | 2 |
| 3 | 2312 | 1200 | 11 | 2322 | 1211 | 14 | 2321 | 1211 | 13 | 3 |
| 4 | 1011 | 2321 | 11 | 1012 | 2221 | 11 | 1011 | 2222 | 11 | 4 |
| 5 | 1212 | 2232 | 15 | 1213 | 2232 | 16 | 1212 | 2231 | 14 | 5 |
| 6 | 1212 | 4354 | 22 | 1222 | 3354 | 22 | 0212 | 3576 | 26 | 6 |
| 7 | 2322 | 3166 | 25 | 2322 | 3156 | 24 | 2322 | 3256 | 25 | 7 |
| 8 | 3433 | 2111 | 18 | 3433 | 2212 | 20 | 4433 | 2201 | 19 | 8 |
| 9 | 1333 | 4343 | 24 | 2333 | 3243 | 23 | 2323 | 4343 | 24 | 9 |
| 10 | 2212 | 2243 | 18 | 2221 | 1243 | 17 | 3211 | 1243 | 17 | 10 |
| 11 | 0021 | 2333 | 14 | 1112 | 2333 | 16 | 0011 | 2233 | 12 | 11 |
| 12 | 3322 | 2344 | 23 | 3222 | 2243 | 20 | 3212 | 2345 | 22 | 12 |
| 13 | 4332 | 2010 | 15 | 4332 | 3111 | 18 | 4332 | 2111 | 17 | 13 |
| 14 | 0013 | 3211 | 11 | 0013 | 2212 | 11 | 0012 | 2112 | 9 | 14 |
| 15 | 3222 | 5555 | 29 | 3222 | 5566 | 31 | 3223 | 6677 | 36 | 15 |
| 16 | 4543 | 3123 | 25 | 4433 | 3213 | 23 | 6653 | 3213 | 29 | 16 |
| 17 | 0121 | 2122 | 11 | 1132 | 2222 | 15 | 1123 | 2122 | 14 | 17 |
| 18 | 3333 | 3343 | 25 | 3334 | 2343 | 25 | 3333 | 3444 | 27 | 18 |
| 19 | 3323 | 3334 | 24 | 3333 | 2334 | 24 | 4322 | 2334 | 23 | 19 |
| 20 | 4333 | 2254 | 26 | 3333 | 2254 | 25 | 3333 | 2254 | 25 | 20 |
| 21 | 4222 | 3455 | 27 | 4222 | 3445 | 26 | 5322 | 2444 | 26 | 21 |
| 22 | 3333 | 2334 | 24 | 3333 | 2324 | 23 | 3332 | 2334 | 23 | 22 |
| 23 | 3222 | 3122 | 17 | 3323 | 3232 | 21 | 3222 | 3222 | 18 | 23 |
| 24 | 2212 | 3434 | 21 | 2212 | 3335 | 21 | 2211 | 3334 | 19 | 24 |
| 25 | 3211 | 1132 | 14 | 2222 | 2132 | 16 | 2211 | 2022 | 12 | 25 |
| 26 | 4332 | 2222 | 20 | 3332 | 2222 | 19 | 3332 | 2111 | 16 | 26 |
| 27 | 2344 | 4665 | 34 | 2334 | 4565 | 32 | 1333 | 4786 | 35 | 27 |
| 28 | 4234 | 3342 | 25 | 4224 | 3342 | 24 | 4223 | 3342 | 23 | 28 |
| 29 | 4323 | 2422 | 22 | 3332 | 2422 | 21 | 3322 | 2532 | 22 | 29 |
| 30 | 2331 | 2443 | 22 | 2332 | 2443 | 23 | 2221 | 2443 | 20 | 30 |
| 31 | 1232 | 3332 | 19 | 1232 | 3342 | 20 | 1222 | 2332 | 17 | 31 |

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, p. 228. Pergamon Press Ltd. Printed in Northern Ireland

## Papers to be published in future issues

G. F. Fooks and I. L. Jones: Correlation analysis of the fading of radio waves reflected vertically from the ionosphere

A. J. Lyon, N. J. Skinner and R. W. Wright: Equatorial spread-$F$ at Ibadan, Nigeria

J. Harwood: Some observations of the occurrence and movement of sporadic-$E$ ionization

F. Horner: Narrow-band atmospherics from two local thunderstorms

H. Rishbeth and C. S. G. K. Setty: The $F$-layer at sunrise

H. Rishbeth: Diffusion of ionization in the sunrise $F$-layer

K. S. Raja Rao: On the seasonal variation in lunar and solar geomagnetic tides in the geomagnetic equatorial region

J. E. Titheridge: A new method for the analysis of ionospheric $h'(f)$ records

H. Unz: A solution of the integral equation $h'(f) = \int \mu'(f, f_0)\, dz(f_0)$

V. I. Krassovsky, N. N. Shefov and V. I. Yarin: On the OH airglow

J. W. King: Magnetic effects in the $F$-region of the ionosphere

J. W. King: Day-to-day and station-to-station correlation of ionospheric $F$-region critical frequencies

J. M. Malville: Excitation of helium in the aurora

O. Awe: The fading of radio waves reflected from the ionosphere at oblique incidence

O. Awe: The fading of radio waves weakly scattered at vertical incidence from heights near 90 km

B. P. Sandford: Aurora and airglow from colour film observations

B. P. Sandford: The behaviour of night-sky 6300, 5577 and 3914 Å emissions at Scott Base, Antarctica

Y. S. N. Murty and S. R. Khastgir: General expression for the group refractive index of the ionosphere

Sir Edward Appleton and A. J. Lyon: Studies of the $E$-layer of the ionosphere—II. Electromagnetic perturbations and other anomolies

H. J. A. Chivers: A statistical study of ionospheric drifts measured by the radio star scintillation technique

H. Unz: On vertical drift velocities of the $F2$-layer

U. Haug: Zur Statistik der räumlichen Helligkeitsschwankungen der [OI]-Linie 5577 Å im Nachthimmelsleuchten

R. Naismith, H. C. Bevan and P. A. Smith: A long term variation in the relation of sunspot numbers to $E$-region character figures.

VOL. 20 1961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 229 to 242. Pergamon Press Ltd. Printed in Northern Ireland

# Correlation analysis of the fading of radio waves reflected vertically from the ionosphere

G. F. Fooks* and I. L. Jones†
Cavendish Laboratory, Cambridge

(*Received* 12 *October* 1960)

**Abstract**—The correlation analysis of Briggs *et al.* (1950) and Phillips and Spencer (1955) has been applied to the fading of radio waves reflected from the $E$- and $F$-regions of the ionosphere, observed at three closely spaced receivers.

Histograms are presented showing the size of the characteristic ellipse of the radio diffraction pattern formed on the ground, and also the axial ratio and the direction of the major axis of this ellipse. Results are presented which show the importance of random changes in the pattern relative to pure drift in producing fading. Corresponding results from other parts of the world are reviewed briefly. Drift velocities found by correlation analysis are compared with those found by the simple time-delay method, and it is shown that the errors in the time-delay method are considerable.

A comparison is made between the method of calculation used by Briggs *et al.* and the six-point method of Yerg (1955); the former is to be preferred.

## 1. Introduction

When a radio wave is reflected from an irregular ionospheric layer, an irregular diffraction pattern is formed on the ground. Information about the spatial properties and movements of this pattern can be obtained by means of observations made on the ground at a small number of points. Usually the fluctuations of the amplitude of the reflected wave are recorded at three points at the corners of a right-angled triangle, and Briggs *et al.* (1950) and Phillips and Spencer (1955) have shown how certain parameters which describe the diffraction pattern can be deduced from such observations. The first step is to calculate the auto-correlation functions of the three fading records, and the cross-correlation functions between the records taken in pairs. It is then possible, by further calculations, to obtain the values of the following quantities:

(a) The size, axial ratio and orientation of a "characteristic ellipse", the radius of which in any direction gives the separation in that direction between two points having a correlation of 0·5. This ellipse describes the average size, degree of elongation, and orientation of the irregularities in the diffraction pattern. If the ellipse is a circle, the correlation is independent of direction and the pattern is said to be isometric.

(b) The magnitude $V$ and direction $\phi$ of the mean velocity of the diffraction pattern over the ground. The angle $\phi$ is measured clockwise from geographic north.

(c) A parameter $V_c$, which has the dimensions of velocity, and which is a measure of the rate at which the pattern changes as it moves. If $V_c = 0$, the pattern drifts without change of form.

The ultimate objective of experiments of this type is to obtain information about the irregularities in the ionosphere. It should be emphasized, however, that

* Now at D.S.I.R. Radio Research Station, Slough, Bucks.
† Now at Royal Aircraft Establishment, Farnborough, Hants.

the quantities which are directly deduced refer to the amplitude pattern on the ground, and in general they are related rather indirectly, through a diffraction process, to the ionospheric irregularities themselves (Ratcliffe, 1956). Thus, the sizes of the ionospheric irregularities need not be even approximately the same as the size of the characteristic ellipse, since if the wave is reflected in such a way that phase variations greater than 1 rad are imposed across its wavefront, the irregularities in the diffraction pattern are smaller than those in the diffracting screen (Hewish, 1951). However, it is expected that the degree of elongation and the orientation of the ionospheric irregularities will be reproduced in the axial ratio and orientation of the characteristic ellipse of the diffraction pattern. These quantities are therefore of particular interest, since they are directly related to the ionospheric irregularities. In addition, it can be shown that the velocity of drift of the ionospheric irregularities is half the velocity of the diffraction pattern over the ground. The parameter $V_c$ has no simple interpretation in terms of the ionospheric irregularities, except in the special case $V_c = 0$, when it may be inferred that the irregularities drift bodily without relative movements.

The parameters which describe the diffraction pattern on the ground have been determined for waves reflected from the $E$- and $F$-regions at Cambridge (52°N, 0°E). Details of the experimental arrangement used to obtain the fading records, the types of record chosen for analysis, and the calculation of the correlation functions, are given in Section 2. In Section 3 the main results are presented, and they are compared with those of other workers in Section 4. In Section 5 certain errors are discussed.

Yerg (1955) has described another method of calculating the parameters of the diffraction pattern. In Section 6, Yerg's method is compared with that of Briggs *et al.* The latter we shall call the "correlation function method" and that of Yerg, the "six-point method".

It is impracticable to use correlation analysis as a routine method of measuring ionospheric drifts, because of the very large number of amplitude readings required from each record. For routine drift measurements, one method which has been widely used, also based on fading records from three closely-spaced receivers, is to determine the mean time delays between corresponding maxima and minima of the fading curves and to deduce a velocity from these (Krautkrämer, 1950; Mitra, 1949). If $\bar{\tau}_x$ is the mean time difference between receivers A and B (Fig. 1), where A is a distance $\xi_0$ east of B, and $\bar{\tau}_y$ is the mean time difference between C and B, where C is a distance $\eta_0$ north of B, then the magnitude and direction of the drift velocity of the diffraction pattern are usually determined from the equations

$$\left.\begin{aligned} V_a &= \xi_0\eta_0(\xi_0^2\bar{\tau}_y^2 + \eta_0^2\bar{\tau}_x^2)^{-\frac{1}{2}} \\ \phi_a &= \tan^{-1}\frac{\eta_0\bar{\tau}_x}{\xi_0\bar{\tau}_y} \end{aligned}\right\} \qquad (1)$$

We shall call $V_a$ and $\phi_a$ the apparent velocity and direction of drift. It can be shown that $V_a = V$ and $\phi_a = \phi$ if the pattern drifts without change of form ($V_c = 0$) and is isometric. This method, which we shall call the time-delay method, gives the correct drift velocity only when these conditions are satisfied. It is



VOL. 20 1961

therefore important to consider whether the properties of the diffraction pattern are such that the time-delay method can be used without too large an error. The observations made at Cambridge are examined from this point of view in Section

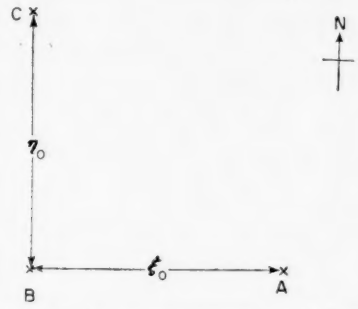


Fig. 1. Example of receiver arrangement for drift measurements.

7. This type of investigation should be made separately for every station at which drifts are measured by the time-delay method, since the properties of the pattern vary from one place to another, as is shown in Section 4.

## 2. Apparatus, Records and Correlation Functions

Pulses from a local transmitter in the frequency range 2–4·5 Mc/s were reflected by the ionosphere at vertical incidence, and were received on three crossed-loop aerials, arranged to accept only the "ordinary" circularly polarized component of the downcoming wave (Phillips, 1951). The aerials were placed at the corners of a triangle which was approximately right-angled; the configuration was varied slightly from time to time, but the spacings were always in the range 100–130 m. Three similar receivers were used, together with a strobing arrangement to select the required echo; the variations in the amplitude of the echo were recorded on a three-track pen recorder.

The records chosen for analysis contained fading that appeared to be statistically stationary, and were twenty or more fading cycles in length. Most of the records had deep fading with more than four fades per minute, though in a few cases the fading was slower or shallower. Between 240 and 360 equally spaced ordinates were measured on each of the three tracks. The correlation functions were calculated on an electronic digital computer for time displacements $\tau$ up to $\pm 20$ times the ordinate spacing. Examples of such auto- and cross-correlation functions are shown in Fig. 2.

The graphical method of Phillips and Spencer (1955) was used in subsequent calculations.

## 3. Results

In all, 111 records were analysed. Of these, thirty-nine were for waves reflected from the $E$-region by day, on frequencies in the range 2·0–2·7 Mc/s, thirty-seven were for waves reflected from the $F$-region by day in the range 3·4–4·5 Mc/s, and thirty-five were for waves reflected from the $F$-region by night in the range 2·0–4·0

Mc/s. The majority of the records were obtained during a few days in April and May 1955 and in February 1958; the remainder were obtained at different times up to May 1960. Magnetic conditions were not disturbed. Results for waves reflected from $E_s$ ionization are not considered in the present paper.

Histograms for the length of the semi-minor axis of the characteristic ellipse are shown in Fig. 3. For waves reflected from the $F$-region, the values tend to be smaller at night than during the day.

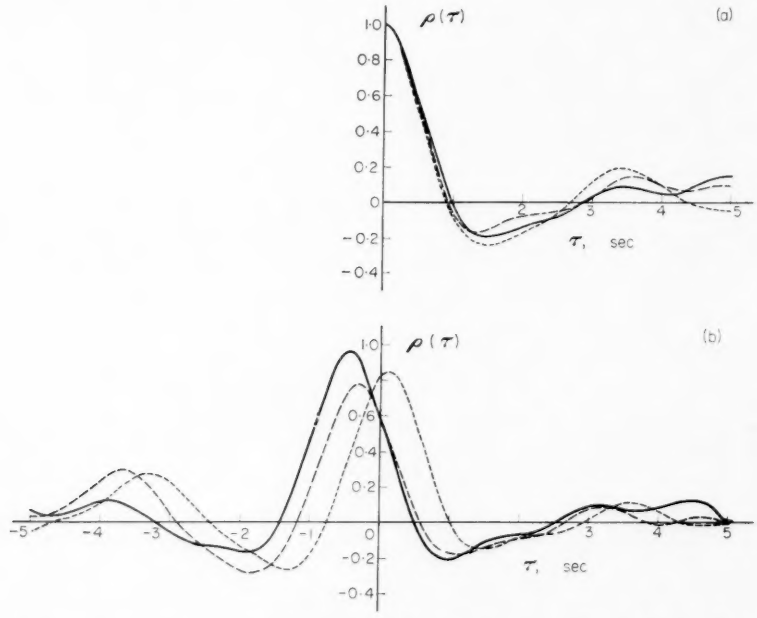


Fig. 2. Correlation functions $\rho(\tau)$ calculated from fading records. 2·4 Mc/s, $E$-region, 19 February 1958, 1430 U.T.
(a) Auto-correlation functions for three receivers.
(b) Cross-correlation functions for three receivers taken in pairs.

Fig. 4 shows histograms for the axial ratio of the characteristic ellipse. The elongation is not great, the median value for each group of results ($E$-region day, $F$-region day, $F$-region night) being less than 2·0. The results for the $F$-region show that the median value of the axial ratio is slightly greater by night (1·8) than by day (1·6).

Fig. 5 shows histograms for the direction of the major axis of the characteristic ellipse. In each group there is a tendency for the major axis to lie in the NW–SE quadrants, and this effect is most marked for the case of waves reflected from the $F$-region by night.

The parameter $V_c$ as defined by Briggs *et al.* depends upon the correlation between the amplitudes observed at two points on the ground. If the pattern is anisometric, this correlation, and therefore the value of $V_c$, will be a function of the direction of the line joining the two points. We use the notation $(V_c)_V$ for the



VOL. 20 1961

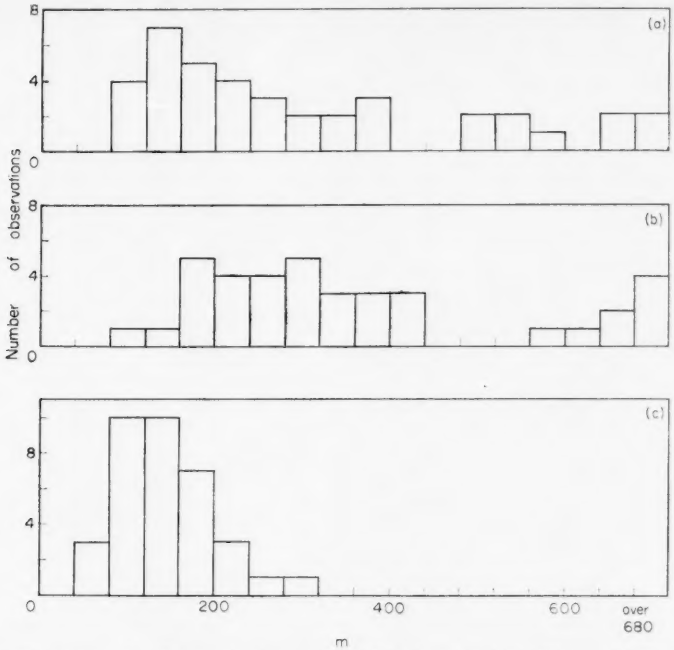


Fig. 3. Histograms for the semi-minor axis of the characteristic ellipse. (a) *E*-region day, median = 235 m. (b) *F*-region day, median = 296 m. (c) *F*-region night, median = 140 m.

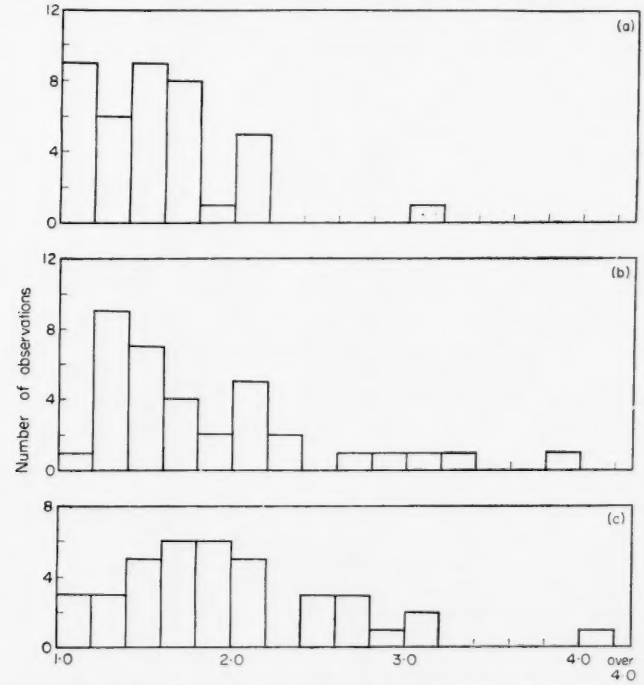


Fig. 4. Histograms for the axial ratio of the characteristic ellipse. (a) *E*-region day, median = 1·5. (b) *F*-region day, median = 1·6. (c) *F*-region night, median = 1·8.

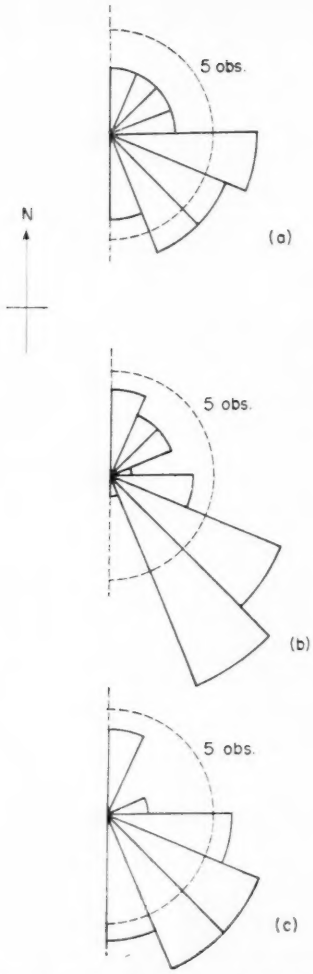


Fig. 5. Polar histograms for the direction of the major axis of the characteristic ellipse. (a) *E*-region day. (b) *F*-region day. (c) *F*-region night.

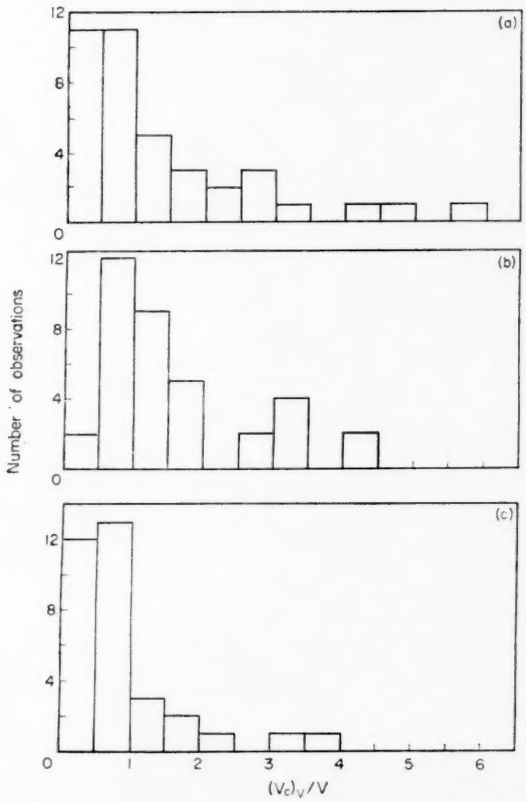


Fig. 6. Histograms for the ratio $(V_c)_V/V$. (a) *E*-region day, median = 0·8. (b) *F*-region day, median = 1·2. (c) *F*-region night, median = 0·7.

value of $V_c$ when this line is in the direction of $V$. It can easily be shown from the analysis of Phillips and Spencer that

$$\frac{(V_c)_V}{V} = \left[\frac{(V_c')_V{}^2}{V^2} - 1\right]^{\frac{1}{2}} \qquad (2)$$

where $(V_c')_V$ is the fading velocity as defined by Briggs *et al.*, also referred to the direction of $V$. The ratio $(V_c)_V/V$ is a convenient measure of the importance of random changes, relative to pure drift, in producing fading.

Histograms for the value of $(V_c)_V/V$ are given in Fig. 6. They show that $(V_c)_V/V$ was less than 1·0 on 56 per cent of occasions for the *E*-region, on 45 per



VOL. 20 1961

cent of occasions for the $F$-region by day, and on 71 per cent of occasions for the $F$-region by night.

## 4. Comparison with Other Workers

Briggs and Phillips (1950), working at Cambridge, used a simple approximate method of estimating the sizes of irregularities in the amplitude pattern on the ground, and obtained results in general agreement with those given in this paper.

Phillips and Spencer (1955), also working at Cambridge, analysed fifteen records using the correlation function method, and found no evidence of very elongated patterns or of a consistent direction of elongation.

Becken and Maehlum (1960) obtained results for the $F$-region at Kjeller (60°N, 11·1°E) similar to those for Cambridge. They found the mean value of the axial ratio to be 1·9, and the most frequent direction of elongation to be SSE, with no evidence of a diurnal variation in either quantity.

Gusev *et al.* (1959), observing at Moscow (56°N, 38°E), used the correlation function method on a small number of records for the $F$-region, and found a mean value of the axial ratio of 2·6, which indicates a slightly greater elongation than at Cambridge. The major axis was in the NE–SW quadrants in every case.

Yerg (1956), working at Puerto Rico (18°N, 67°W) and using the six-point method, found axial ratios of the order of 1·8 for waves reflected from the $E$- and $F1$-regions. The direction of elongation was most frequently in the NE–SW quadrants. Yerg also found values of $(V_c)_V/V$ (which he called $V_i/V$) of the order of 2·5, which is somewhat greater than is observed at Cambridge.

In these observations, all made in medium latitudes, the axial ratio is found to be small, usually less than 2·0, and there is a rather ill-defined distribution of directions of elongation.

In contrast, Skinner *et al.* (1958) working at Ibadan, near the magnetic equator ($2\frac{1}{2}$°S magnetic latitude), found very elongated amplitude patterns for waves reflected from both $E$- and $F$-regions. The direction of elongation was parallel to the earth's magnetic field.

## 5. Errors in Correlation Analysis

There are two distinct sources of error in correlation analysis. First, there are random errors in the correlation functions which arise because the records from which they are derived are limited in length. In addition, there will be further errors if the records are not statistically stationary. If, for instance, there is a progressive change in the velocity of drift during the course of a record, the overall values of the velocity and of the ratio $(V_c)_V/V$ for the record will be misleading.

Random errors in the auto-correlation functions lead to small differences between the functions for the three receivers. These differences can be seen in Fig. 2(a). The standard error $\sigma$ of a correlation coefficient $\rho$, derived from a large number $N$ of independent pairs of values, is given by

$$\sigma = \frac{1 - \rho^2}{N^{\frac{1}{2}}} \qquad (3)$$

if $\rho$ is small.

This formula cannot be applied directly because the original fading curves are auto-correlated and successive ordinates are not independent of each other. It is therefore not clear what value of $N$ should be used in equation (3) to calculate the random error. Also, since the fading curves are highly correlated with each other, they are not independent samples of the fading and so the differences between the auto-correlation functions do not give a true indication of the random error. If, however, two consecutive records showing apparently similar fading, and approximately equal in length, are analysed separately, the two resulting sets of auto-correlation functions will have been derived from independent samples of the fading. We get the best available value of $\rho(\tau)$ for each set by taking the mean of the functions for the three receivers; this is a valid procedure if the differences are small. It is not possible to calculate the standard error from only two mean values, but the difference between the mean functions can be compared with the differences

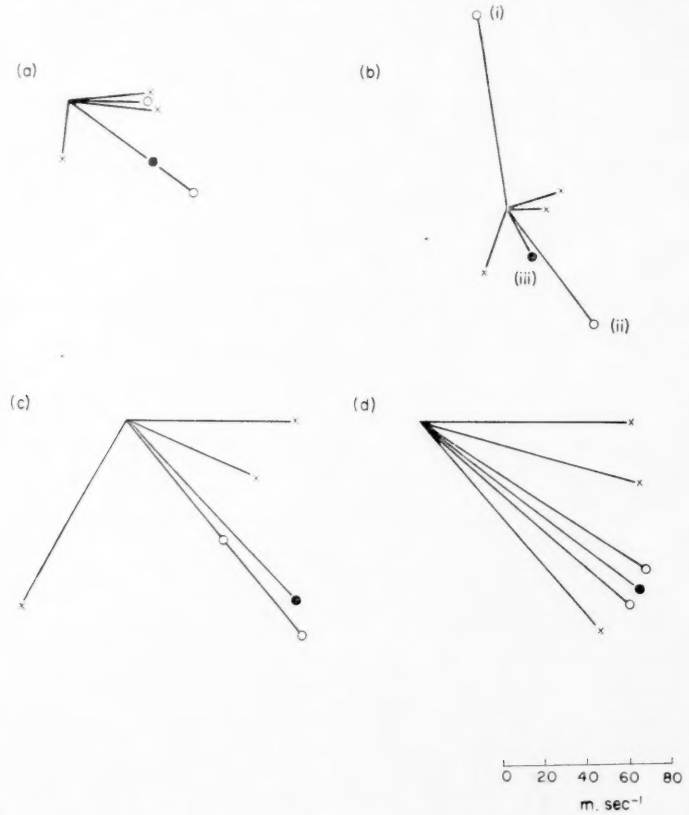


Fig. 7. Velocity vectors deduced on four occasions using the six-point method with different values of $\tau_i$, and using the correlation function method. 2·01 Mc/s, $E$-region.
—— ×, Six-point method (three values of $\tau_i$); —— ●, Correlation function method, combined records; —— ○, Correlation function method, consecutive records.
(a) 28 April 1955, 0800 U.T. $(V_c)_F/V = 2{\cdot}2$.
(b) 28 April 1955, 0932 U.T. (i) $(V_c)_F/V = 1{\cdot}4$. (ii) $(V_c)_F/V = 1{\cdot}0$. (iii) $(V_c)_F/V = 4{\cdot}7$.
(c) 29 April 1955, 1205 U.T. $(V_c)_F/V = 0{\cdot}4$.
(d) 29 April 1955, 1335 U.T. $(V_c)_F/V = 0{\cdot}5$.



VOL. 20 1961

between the individual functions. If this test is performed on a number of occasions, it is possible to get an indication of the true importance of random errors. The parameters of the ground diffraction pattern can also be calculated separately for two consecutive records.

The fading observed on six occasions was analysed in this way. In the part of the function used in subsequent calculations, i.e. over the range from $\rho(\tau) = 1$(at $\tau = 0$) down to $\rho(\tau) = 0{\cdot}2$, the differences between the mean auto-correlation functions were comparable with the spread in the values of the individual functions in all cases except one. This test shows that the random errors are not usually much greater than the differences between the individual auto-correlation functions for the three receivers. It also shows that it is possible to select for analysis portions of record in which the fading is sufficiently uniform to give consistent auto-correlation functions, and which consequently can be regarded as statistically stationary. If the fading had been non-stationary, much greater differences between the mean auto-correlation functions would have been expected.

Fig. 7 shows values of the drift velocity found by different methods on four occasions. The vectors marked by circles were found by the correlation function method from consecutive records. Those marked by dots were derived from the two records combined into a single length of record, and the values of $(V_c)_V/V$ in Fig. 7(a), (c) and (d) were also derived from the combined records. (The vectors marked by crosses were obtained by the six-point method; these will be discussed in the next section). The differences between the velocities for the consecutive records were greater when $(V_c)_V/V$ was large. In case (b), the large difference between the velocities is probably due to a real change in the velocity during the course of the observation; the values of $(V_c)_V/V$ for the records separately are both less than the value of $(V_c)_V/V$ for the combined record.

The results in Fig. 7 show that there can be considerable random errors in the drift velocity found by correlation analysis, perhaps as much as $\pm 50$ per cent in unfavourable cases. The errors in the other parameters of the ground diffraction pattern are of the same order. In particular, the value of $(V_c)_V/V$ is likely to be inaccurate, as we now show.

$(V_c)_V/V$ is calculated from equation (2). When differentiated, equation (2) gives

$$\frac{d[(V_c)_V/V]}{d[(V_c')_V/V]} = \frac{(V_c')_V/V}{(V_c)_V/V}. \tag{4}$$

Therefore, an error in $(V_c')_V/V$ produces a large error in $(V_c)_V/V$ when $(V_c)_V/V$ is small; in this case $(V_c')_V/V \simeq 1$.

The analysis makes use of a quantity $V'$, which is in effect the apparent velocity $V_a$ corrected for elongation of the diffraction pattern. $V'$ is found from the time displacements required to give maximum cross-correlation between the fading records.

When $(V_c)_V/V$ is large,

$$V' \geqslant (V_c')_V \geqslant V.$$

In such cases, $V'$ is large, and must be deduced from time delays much smaller than the period of fading.

From equation (2)

$$V^2 + (V_c)_V{}^2 = (V_c')_V{}^2$$

$$\therefore \text{ if } (V_c)_V \gg V$$

$$(V_c)_V \simeq (V_c')_V$$

and an error in $V'$ will make little difference to the calculated value of $(V_c)_V$.

Therefore, we may say that $V$ is found more accurately than $(V_c)_V$ when $(V_c)_V/V$ is small, and vice versa when $(V_c)_V/V$ is large. In either case, the accuracy of $(V_c)_V/V$ itself is poor, and it serves as little more than an indication of the main cause of the fading.

All the parameters found by correlation analysis show great variation from time to time, so provided a sufficiently large number of results is obtained, the frequency distributions for the various quantities should not be greatly in error, in spite of the inaccuracy of individual results.

## 6. Comparison of the Correlation Function Method with the Six-point Method

Yerg (1955) pointed out that an auto-correlation function in two space dimensions and in time can be represented as a Taylor's series in three variables, and that near the origin only the first- and second-order terms need be retained.

Thus he wrote

$$\rho(\xi,\eta,\tau) = 1 + A\xi^2 + B\tau^2 + C\eta^2 + 2H\xi\tau + 2M\eta\tau + 2N\xi\eta \qquad (5)$$

the coefficients of the first-order terms being zero because the function must be symmetrical about the origin. The six coefficients can be evaluated from experimental data. Suitable data are:

(i) the three values of the instantaneous cross-correlation between the three pairs of fading records,

(ii) the mean value of the auto-correlation at some time displacement $\tau_i$, and

(iii) two of the three values of the cross-correlation at the same time displacement $\tau_i$.

The six coefficients $A, B, C, H, M, N$ in (5) can then be calculated. When these coefficients have been determined, it is possible to calculate the same parameters of the diffraction pattern as are found by the correlation function method.

Ideally, the six-point method and the correlation function method should give identical results. The computed values of correlation are subject to errors, however, and bearing this in mind we wish to consider which method of calculation is *a priori* more reliable. We shall then examine the results of a comparison between the methods.

Even though each point on a correlation function $\rho(\tau)$ is to some extent inaccurate, the function itself is a smooth curve (Fig. 2). Therefore the error changes by a small amount between successive calculated points of the function, and over a small part of the curve, changes in $\rho$ will be given more accurately than $\rho$ itself.



VOL. 20 1961

Also, the position in time of the maximum cross-correlation will be found more accurately than the value of $\tau$ corresponding to any specific value of $\rho$. It would seem better to make use of the position of maximum cross-correlation, as is done in the correlation function method, than to base the calculation entirely on values of $\rho$ as in the six-point method.

Four records were analysed by the six-point method. The coefficients of the analytic function were modified to allow for the fact that the three aerials were not exactly at the corners of a right-angled triangle. The analysis was done three times in each case, using different values of $\tau_i$ near the origin and such that only values of correlation greater than 0·3 were used. In most cases, the values were greater than 0·5 and were therefore taken from the most accurate parts of the correlation functions.

The occasions chosen were the same as those for which consecutive records were analysed separately by the correlation function method, as described in Section 5. The six-point method was applied to the combined records. The values of drift velocity obtained are shown in Fig. 7, the velocity vectors being marked by crosses. They show considerable variation, according to which value of $\tau_i$ is used, even when $(V_c)_V/V < 1$.

Sales (1956) applied the six-point method to two fading records on 75 kc/s. He also found large variations in the parameters as the value of $\tau_i$ was changed.

In using the correlation function method, it is necessary to calculate a much larger number of correlation coefficients than is needed in the six-point method. This is not a practical difficulty if the coefficients can be found with the aid of an electronic computer. On the whole we consider that the correlation function method makes better use of the available data.

## 7. Comparison with the Time-delay Method of Drift Measurement

Insufficient results have been obtained by correlation analysis to make it possible to look for any mean daily or seasonal variations in the velocity of drift, using these results alone. However, the results which are available can be compared with the apparent velocities obtained by the time-delay method on the same occasions.

The apparent velocity $V_a$ is found by measuring the time differences between corresponding maxima and minima of the fading curves. The records that are chosen for this type of analysis are those in which there is obvious similarity between the fading curves and in which the time differences are consistent. Of the 111 records to which correlation analysis was applied, fifty-four satisfied the above criteria, and in these cases the apparent velocity was determined.

Histograms for the ratio $V_a/V$ are shown in Fig. 8. The results show that $V_a$ is greater than $V$ in the majority of cases, so that even when the fading records appear to be suitable for analysis by the time-delay method, the deduced magnitudes of the drift velocity may be considerably too great. For all the results taken together, the average value of $V_a/V$ is about 1·4.

The differences $(\phi_a - \phi)$ between the apparent and true directions of drift are

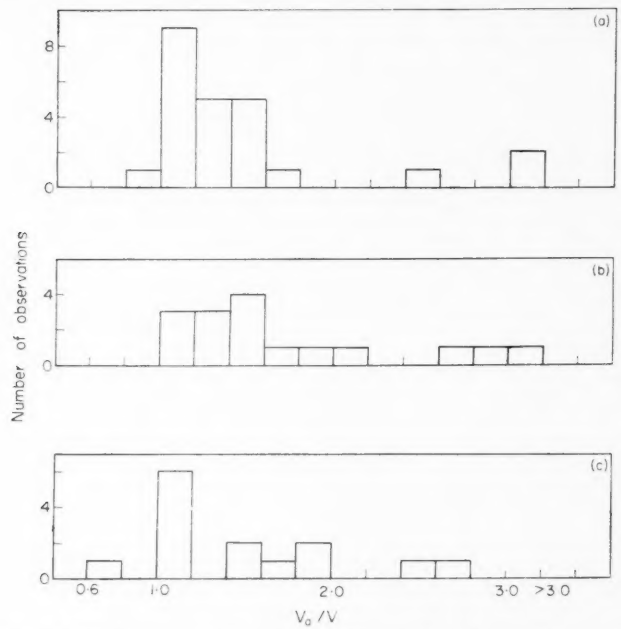


Fig. 8. Histograms for the ratio $V_a/V$. (a) $E$-region, median = 1·4. (b) $F$-region day, median = 1·5. (c) $F$-region night, median = 1·3.

plotted in Fig. 9. There is no systematic difference for the $E$-region or for the $F$-region by day. At night, when only reflections from the $F$-region were observed, most of the results give $\phi \simeq 280°$ and in these cases the value of $(\phi_a - \phi)$ is systematically negative, the mean being $-18°$. This result is consistent with the observed distribution of directions of elongation of the diffraction pattern for waves reflected from the $F$-region by night (Section 4).

Yerg (1956) made a similar comparison, using the six-point method, for his observations in Puerto Rico. The results showed that the mean apparent velocity was three to five times as great as the mean true velocity.

## 8. Conclusions

The correlation analysis of the fading of radio waves reflected from the $E$- and $F$-regions at Cambridge (52°N, 0°E) shows that the elongation of the diffraction pattern on the ground is not great, but there is a tendency for the direction of elongation to be in the NW–SE quadrants, particularly for waves reflected from the $F$-region at night. For the $F$-region, the scale of the pattern is smaller at night than during the day.

The ratio $(V_c)_V/V$, which describes the importance of random changes relative to drift, was less than 1·0 for about half the cases observed.

Apparent drift velocities found by the time-delay method are usually too great. There is a wide range of errors; on the average, the apparent velocity is about 1·4 times the true velocity. A systematic error in the apparent direction of drift was



VOL. 20 1961

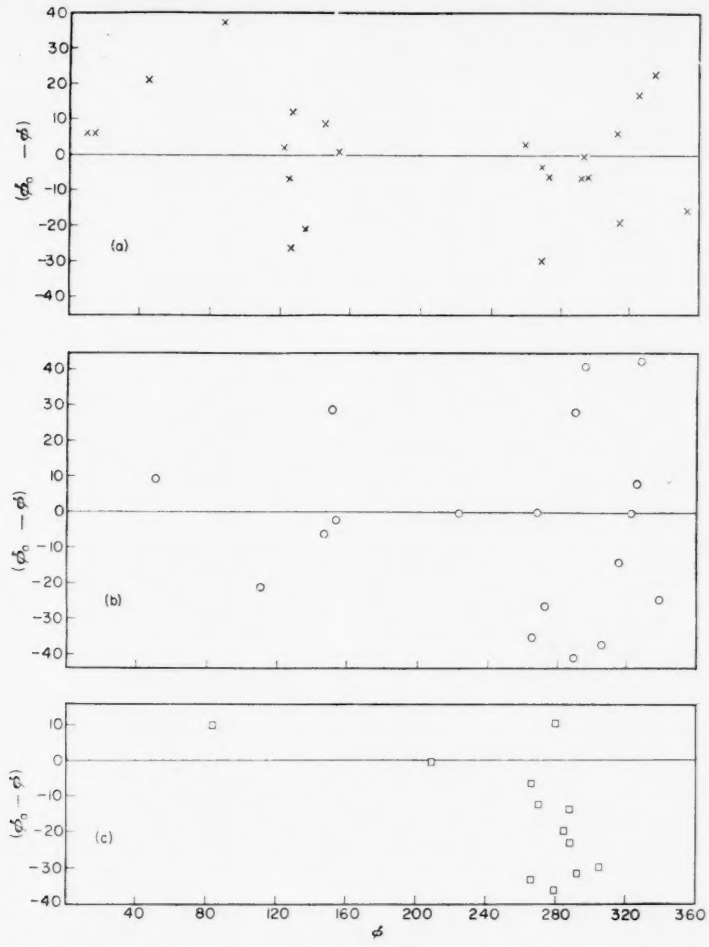


Fig. 9. Differences ($\phi_a - \phi$) between the apparent and true directions of drift, plotted against the true direction of drift $\phi$. (a) $E$-region. (b) $F$-region day. (c) $F$-region night.

found only for the $F$-region at night, the mean error being $-18°$ for true drift directions $\phi$ of about 280°.

The above conclusions apply to occasions when the fading is deep and reasonably fast. At other times, there is a tendency for the sizes of the characteristic ellipses to be large; this makes the analysis difficult to apply. Such occasions are not used in drift measurements.

These results obtained at Cambridge are in quite good agreement with results obtained at other places in medium latitudes.

The analysis of the records by the correlation function method makes better use of the available data than the use of the six-point method, and is to be preferred if the complete correlation functions can be computed.

*Acknowledgements*—We are indebted to the Department of Scientific and Industrial Research (G. F. F.) and to the Gassiot Committee of the Royal Society (I. L. J) for the award of maintenance grants. The work forms part of a programme for the investigation of ionospheric irregularities, supported by a grant from the Department of Scientific and Industrial Research.

We wish to thank the Director of the Mathematical Laboratory, University of Cambridge, for permission to use the EDSAC computers.

Much tedious work has been done in reading records and plotting correlation functions by the computing staff of the Radio Section of the Cavendish Laboratory, and we wish to express our thanks to them. We have benefited much from discussions with Mr. J. A. Ratcliffe and Dr. B. H. Briggs.

## References


| | | |
|---|---|---|
| Becken J. and Maehlum B. | 1960 | *J. Geophys. Res.* **65,** 1485. |
| Briggs B. H. and Phillips G. J. | 1950 | *Proc. Phys. Soc. Lond.* B **63,** 907. |
| Briggs B. H., Phillips G. J. and Shinn D. H. | 1950 | *Proc. Phys. Soc. Lond.* B **63,** 106. |
| Gusev V. D., Kushnerevsky J. V. and Mirkotan S. F. | 1959 | U.R.S.I.–A.G.I. Symposium, Brussels. |
| Hewish A. | 1951 | *Proc. Roy. Soc.* A **209,** 81. |
| Krautkrämer J. | 1950 | *Arch. Elektr. Übertr.* **4,** 133. |
| Mitra S. N. | 1949 | *Proc. Instn Elect. Engrs* III **96,** 441. |
| Phillips G. J. | 1951 | *Proc. Instn Elect. Engrs* III **98,** 237. |
| Phillips G. J. and Spencer M. | 1955 | *Proc. Phys. Soc. Lond.* B **68,** 481. |
| Ratcliffe J. A. | 1956 | *Rep. Progr. Phys.* **19,** 188. |
| Sales G. S. | 1956 | *Penn. State Univ. Rep.* No. 88. |
| Skinner N. J., Hope J. and Wright R. W. | 1958 | *Nature, Lond.* **102,** 1363. |
| Yerg D. G. | 1955 | *J. Geophys. Res.* **60,** 173. |
| Yerg D. G. | 1956 | *J. Atmosph. Terr. Phys.* **8,** 247. |

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 243 to 262. Pergamon Press Ltd. Printed in Northern Ireland

# Some observations of the occurrence and movement of sporadic-*E* ionization

J. HARWOOD
D.S.I.R. Radio Research Station, Slough, Bucks.*

(*Received* 21 *October* 1960)

**Abstract**—A rotating-aerial back-scatter sounder has been used at 17 Mc/s to examine the location, approximate size, and movements of clouds of sporadic-*E* ionization for the period May–August 1959, as part of a European co-operative experiment involving a number of vertical-incidence sounding stations. Clouds, of average diameter about 200 km and average duration 2 hr, were seen more frequently to the south than to the north. One third of them drifted, predominantly towards the south-west, at a mean speed of 60 m/sec. There was good correlation between the occurrence of back-scatter echoes propagated by way of $E_s$ clouds, one-way propagation over a 700 km path by the same means, and the occurrence of high $E_s$ critical frequencies at a few vertical incidence sounders. To account for a lack of correlation between other drift measurements and $E_s$ cloud movements, it is suggested that the former are based on the fine structure of the cloud, whereas the latter depend on observation of the main cloud boundaries.

## 1. INTRODUCTION

RADIO SOUNDINGS of the ionosphere at vertical incidence show certain irregular echoes from heights near 100 km which do not fit into the regular pattern expected for the normal *E*-layer. These are generally classified as echoes from sporadic-*E* ($E_s$) ionization, which may be envisaged in the form of a thin cloudy layer whose critical frequency may be well above that of the *E*-layer itself. To study the nature of the "cloudiness" in the horizontal plane, results from a number of vertical incidence sounders may be correlated, (BENNINGTON, 1952; MITRA, 1954), and with this objective in mind, a European Co-operative Experiment was organized for the study of $E_s$ in the summer months of 1959. As part of this experiment, the ground-back-scatter technique (SHEARMAN, 1956; PETERSON, 1957) was used at Slough, England, to study clouds of $E_s$ ionization occurring within 700 km of the site.

In this paper, the occurrence and movement of $E_s$ clouds as deduced from the back-scatter observations are examined, and the data are correlated, where possible, with vertical incidence soundings and with ionospheric propagation over a particular 700 km path. In this way, an insight into the nature of $E_s$ ionization is sought.

Comparisons are also made with earlier observations of the characteristics of $E_s$ clouds by the back-scatter technique at Stanford, U.S.A. (PETERSON, 1957) and at Slough during the I.G.Y. (SHEARMAN and HARWOOD, 1960). Another technique which has been used to study the spatial characteristics of $E_s$ ionization is based on observations of v.h.f. propagation via the ionization (GERSON, 1950).

## 2. EXPERIMENTAL METHODS

The back-scatter technique consists in transmitting, from a directional aerial, radio-frequency pulses which irradiate regions of the earth after reflection from the ionosphere; the signals scattered back along the same path from the irradiated ground irregularities are received on the same aerial and displayed on cathode-ray

* Official communication.

tubes. The delay between the transmitted pulses and the received echoes gives a measure of the range of the ground irregularities, and hence of the areas of the ionosphere (at the mid-point of the path) responsible for their irradiation. In the present experiments, a rotating Yagi aerial was used at 17 Mc/s to survey the ground areas to distances of about 2000 km. The time-base of a plan-position indicator (p.p.i.) was rotated in synchronism with the aerial, and the range and azimuth of the irradiated zones of the ground were thus presented in polar form (Fig. 1). A detailed description of the equipment has previously been given (SHEARMAN and MARTIN, 1956). Echoes due to propagation by way of sporadic-*E* ionization were distinguished by their short range, time variation and localized nature, and their study enabled deductions to be made about the occurrence and movement of the ionization itself.

The p.p.i. display was photographed on 16 mm ciné film at 1 min intervals throughout the days of observation. When viewed on a normal ciné projector, the films showed the days' occurrences speeded up about 1000 times. Analysis of the data consisted in measuring, frame by frame, the minimum and maximum range, the azimuthal extent, and the centre azimuth of each echo on the display. Linear and polar graphs were made of these parameters. Echoes were shown as blacked-out portions of the p.p.i. display which contained little information about echo amplitude. To include some such data, and to assist in distinguishing echoes at nearby aximuths which, at high gain, overlapped appreciably, a switched-gain facility was introduced; a threshold bias was applied so that only those parts of the echoes which exceeded a certain amplitude were shown on the display. Three thresholds, at 10 dB intervals, were applied in sequence to successive frames of the display, to enable a coarse resolution of amplitude to be made. Additional amplitude information was obtained from photographs of the normal range–amplitude display of the echoes. These photographs were taken periodically when the aerial was pointing in directions of interest.

VOL. 20 1961

### 2.1. *Propagation over a 700 km path*

Radio-frequency pulses at a frequency of 17 Mc/s were radiated from a directional aerial at Inverness, 704 km from Slough. These pulses were not normally received at Slough, which was in the skip-zone for propagation via *E*- and *F*-layers. Their occasional reception at Slough thus indicated the presence of sporadic ionization.

The repetition rate (1 or 100 per sec) was controlled by a crystal clock at Inverness, and the pulses were observed on a similarly controlled time-base display at Slough. A stable pulse pattern could thus be observed at Slough, and it was possible to distinguish between different modes of propagation by the effect on pulse delay or shape; in particular, propagation by way of $E_s$ ionization at the mid-point of the path could be distinguished from occasional propagation via auroral ionization.

### 2.2. *Vertical-incidence sounders*

Of the many vertical-incidence sounding stations taking part in the co-operative experiment, five were situated below the areas of the ionosphere of interest in the

Fig. 1. Plan position-indicator showing $E_s$ cloud. Range circles at 200 km intervals. Ground scatter received by way of an $E_s$ cloud is indicated by the dark area in the south-west quadrant. The minimum and maximum slant ranges from transmitter to ground by way of the $E_s$ cloud are 590 and 1020 km.

VOL.
20
1961

back-scatter experiments (Slough, De Bilt, Dourbes, Bagneux and Swansea). In addition, a temporary sounder was installed near Kirkby Stephen, Westmorland, within 1 km of the mid-point of the path from Slough to Inverness, during May 1959. The location of these sounders is shown in Fig. 2.

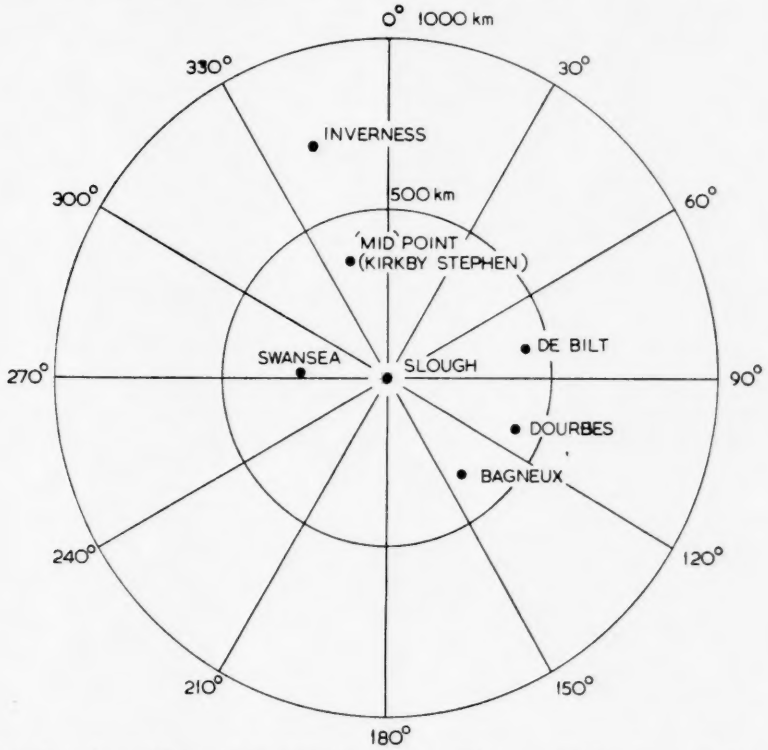


Fig. 2. Locations of European co-operative stations. Distances are slant ranges to $E_s$ layer above each station.

## 3. Statistical Data

### 3.1. *General*

Observations of $E_s$ clouds by the back-scatter technique during the I.G.Y. revealed a few clearly distinguishable features, for example:

(i) Clouds occurred predominantly in the summer daytime, and none were seen in winter.

(ii) There were many more clouds at latitudes south of Slough than at those to the north.

(iii) A mean cloud diameter of 150 km was deduced.

In the 1959 experiments, observations were made on 4 days in each of the months May, June, July and August. The times of occurrence of $E_s$ ionization are shown in Fig. 3. An examination of these observations for the characteristics mentioned above reveals similar conclusions; a mean cloud diameter of 160 km was deduced from measurements of the extent in range of all the clouds observed.

As minimum ranges are affected by skip-distance considerations and maximum ranges by attenuation, the value of 160 km is likely to be too small, and evidence is given later for a mean cloud diameter nearer 250 km.

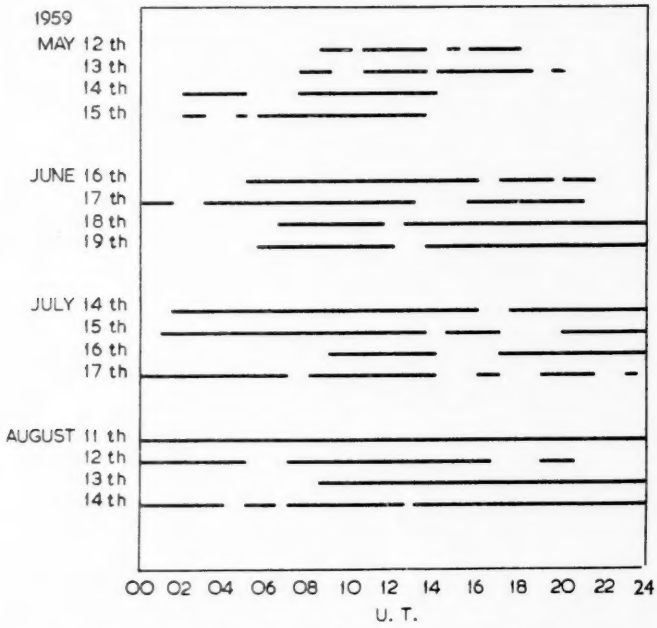


Fig. 3. Times of occurrence of $E_s$ ionization on back-scatter plan position indicator.

3.2. *Azimuthal occurrence of echoes*

Two features of the azimuthal occurrence were investigated. Firstly, an attempt was made to express more quantitatively the predominance of echoes at latitudes south of Slough. Secondly, a search was made for any dependence of the azimuth on time of day, which might be expected in view of the apparent solar dependence of $E_s$ occurrence.

3.2.1. *Variation of occurrence with latitude.* Some world-wide contours of the percentage of time for which the vertical-incidence critical frequency ($f_0E_s$) of sporadic-$E$ ionization exceeds 5 Mc/s have been published (SMITH, 1955). These show that during the summer months this condition would be expected in an area south-east of Slough for about 22 per cent of the time, and north-west for about 15 per cent. Considering ordinary ray propagation, a critical frequency of 5 Mc/s corresponds to a skip-distance at 17 Mc/s of about 800 km, and a measure of the percentage of time $fE_s$ exceeds 5 Mc/s should be given by the percentage of time for which backscatter echoes come from a point 800 km from Slough via $E_s$ ionization. Ambiguities in the consideration of ordinary or extraordinary ray propagation, and the actual height of the $E_s$ layer might introduce errors in range of about 50 km. A plot was therefore made of the occurrence of $E_s$ ionization, as shown by back-scatter echoes from ranges near 800 km. In Fig. 4, the occurrences



VOL. 20 1961

are shown as a function of azimuth, in the form of a histogram representing the number of hours during the 16 days of the experiment for which back-scatter was seen anywhere in the range 700–900 km and in each 10° of azimuth. A 200 km spread in range was chosen in preference to the use of 800 km exactly, as the number of available results could thereby be substantially increased. The azimuthal extent of an echo was determined more by the beam-width of the rotating

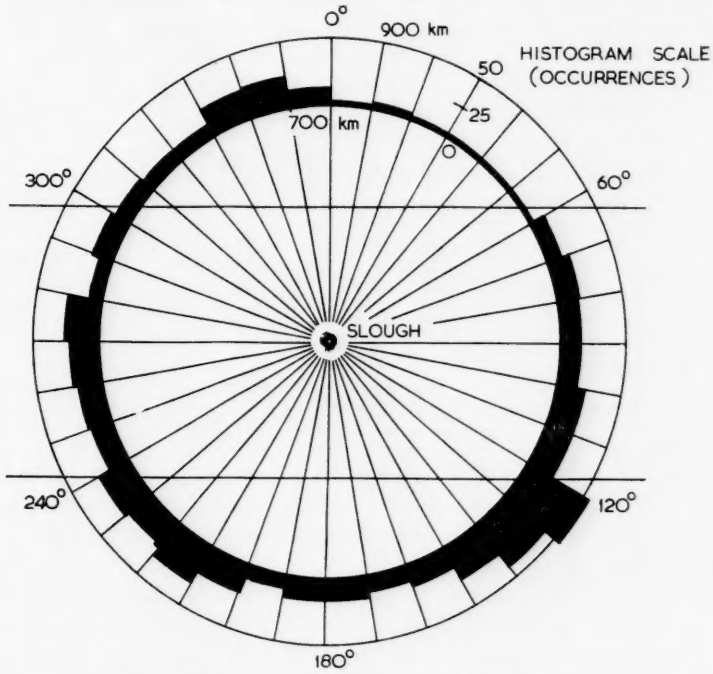


Fig. 4. Half-hourly occurrences of $E_s$ in 10° azimuths.

aerial than by the size of the $E_s$ cloud, and therefore the centre azimuths have been used for the data in the figure. It is shown in the Appendix that, assuming circular clouds with a radius of 100 km, the percentage of time for which ionization ($fE_s > 5$ Mc/s) would be seen overhead by a vertical-incidence sounder at a given point in the range 700–900 km is about 1·4 times the percentage given in Fig. 4 for a 10° azimuthal extent. Using this factor and taking averages of the data, in the regions above, between and below the horizontal lines in Fig. 4, one may deduce the percentages given in Table 1 at latitudes of about 49°, 52° (Slough) and 55°. These results are compared with the corresponding information from vertical-incidence sounders at Slough and Inverness, and with the world contours given by Smith (1955).

A decrease of $E_s$ occurrence with increasing latitude is clearly shown by both vertical-incidence soundings and by back-scatter, but the percentage of time for which back-scatter is seen from any point is very much less than vertical-incidence soundings at the mid-point would indicate. The reason for this is not simply that

the back-scatter system is less sensitive than a vertical sounder; the calculated minimum ground area detectable by back-scatter is very much less than that which would be irradiated by perfectly reflecting clouds of the dimensions indicated above, being of the order of a few square kilometres. The way in which echoes form, described in the next Section, suggests that on several occasions a cloud of ionization already present in the ionosphere becomes detectable by the back-scatter system because of the removal of some form of echo attenuation. The presence of an attenuating mechanism for some of the time would account for the above discrepancies. The explanation is unlikely to be in terms of non-deviative absorption, as neither the fluctuations nor the mean absorption would be of sufficient magnitude. If, however, the clouds had a fine structure, which resulted in diffuse

Table 1. Percentage of time for which $f_0E_s \geqslant 5$ Mc/s during 16 Days in May, June, July, August 1959

| Station | Latitude (deg.) | Vertical-incidence (%) | Back-scatter (%) | World contours (%) |
|---|---|---|---|---|
| Inverness | 57 | 6 | — | ~10 |
| | 55 | — | 2 | 15 |
| Slough | 52 | 14 | 3·3 | 19 |
| | 49 | — | 4·5 | 22 |

reflection, a high attenuation of back-scatter echoes would be expected, particularly in the case of the smaller clouds. Some evidence of this fine structure is provided by vertical incidence soundings.

3.2.2. *Variation of azimuth with time.* The second investigation was to determine whether the azimuth at which echoes occurred had a diurnal variation. Except for an apparent concentration towards the south-east near dawn, statistical plots showed no trend of azimuthal variation with time of day. Echoes towards the north, which might be expected to correlate with auroral behaviour, also showed no diurnal trend.

## 4. Formation and Decay of Echoes

The occurrence of an $E_s$ echo on a vertical-incidence sounder may mean that an $E_s$ cloud has formed, but could also mean that a cloud has drifted overhead. The much larger area surveyed by the back-scatter system should enable these phenomena to be resolved, and a study of the actual formation and decay should be useful in understanding the nature of $E_s$ ionization.

An echo appeared on the p.p.i. display as an arc, which sometimes subsequently intensified (by increase in echo amplitude) or increased in area. A distinction could be drawn between echoes which formed in the "field of view" of the aerial, and those which moved into view from outside the limiting range for detection (or, occasionally, from inside the skip zone). The formation and decay of about forty echoes were examined; of these, twenty-five were in the former category and fifteen in the latter.

4.1. *Formation in field of view*

If a local patch of $E_s$ ionization were to grow from zero in electron density and extent, a point would be reached at which conditions of critical frequency and magnitude were such as to give a back-scatter echo. The maximum and minimum ranges of this echo, defining certain critical frequency contours of the cloud, would subsequently be expected to diverge as the clouds grew larger. Of the twenty-five echoes examined, only a third behaved in this way. The rest appeared with greater extent in range than the minimum observable, as though a cloud of already determined magnitude (about 150 km) gradually became "visible" through some change in its reflecting properties (see Section 3.2). For example, it is possible to envisage a fine structure of "blobs" within the cloud (a structure suggested by many workers, e.g. ECKERSLEY, 1932) which at first reflects diffusely, but which subsequently behaves more like a mirror as the background ionization increases. This change in reflecting properties appears to take place over a large area at one time.

4.2. *Movement into field of view*

Of the forty echoes examined, the formation of about fifteen could be explained in terms of a cloud of ionization drifting into the field of view. Echoes due to $E_s$ propagation were rarely seen at ranges beyond 1400 km, owing to the combination of aerial characteristics, spatial attenuation, and ground scattering properties in attenuating the back-scatter. Thus if an echo appeared near this limiting range, say between 1000 and 1200 km, and subsequently its minimum and maximum range decreased together, the simplest interpretation was in terms of a drifting $E_s$ cloud. Some examples of drifting echoes are shown in Fig. 5; the range extents are shown as measured from the display. Figs. 5(a) and 5(b) show examples of simple drifts, in which the minimum and maximum ranges move at the same speed. Sometimes, however, the minimum range decreased whilst the maximum range showed little change (Fig. 5c). Such a variation was still best interpreted as a moving cloud as distinct from an expanding one, it being assumed that the outer edge of the cloud was beyond the maximum observable range. Fig. 5(d) shows an outward drift, and Fig. 5(c) an example of a cloud drifting inwards so far that it disappears inside the skip zone at a range of 400 km.

Measurements of the mean rate of movement into or out of the field of view in thirty instances led to a mean drift velocity of 390 km/hr or 110 m/sec. The changes of azimuth, which will affect these velocities, have for clarity been omitted from Fig. 5. Their effect will be to make only a small increase in the mean drift velocity, as in most of the examples the rate of change of azimuth was much less significant than that of range.

## 5. Duration and Movement of Echoes

5.1. *Duration*

The period during which a particular echo on the display remains visible may be from a few minutes to several hours. The main difficulty in examining echo duration is that of defining whether, after a period of several minutes break or fading out, the re-appearing echo can be considered as the same one. The results

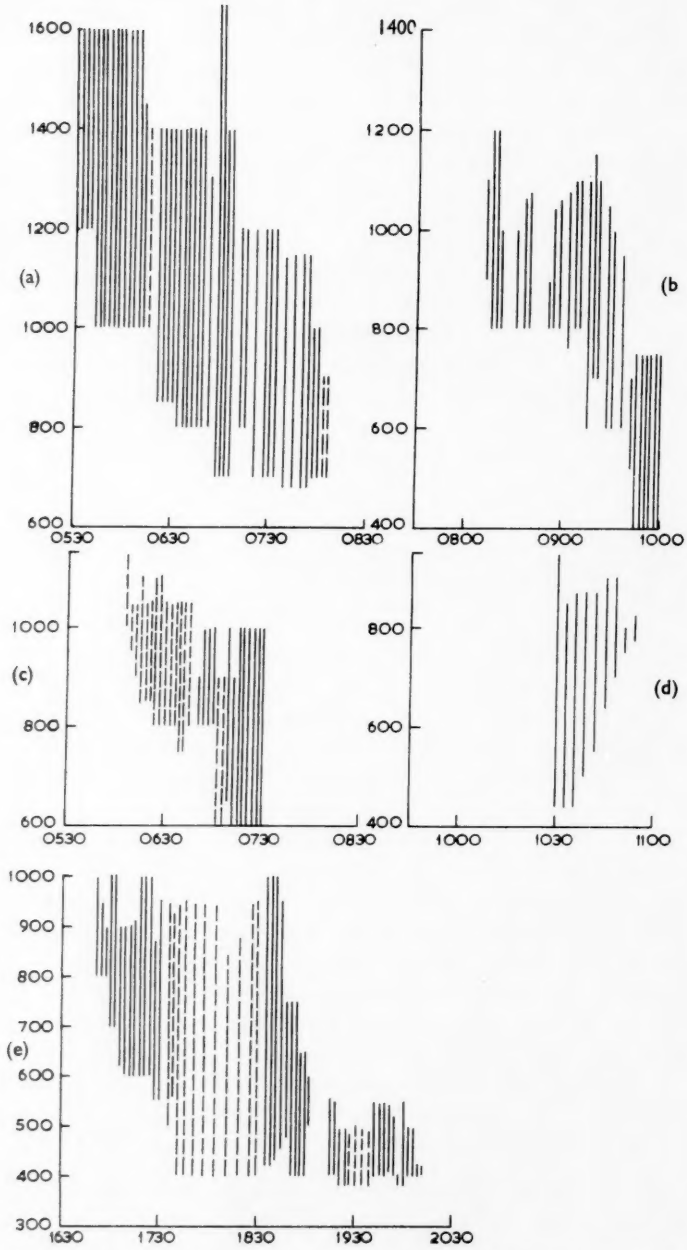


Fig. 5. Examples of echo formation and decay. (a) 19 June 1959, 0530–0800 hours. (b) 14 May 1959, 0800–1000 hours. (c) 15 May 1959, 0600–0715 hours. (d) 15 May 1959, 1030–1100 hours. (e) 17 June 1959, 1650–2010 hours.



VOL.
20
1961

obtained in the summer of 1959 have been used to form the histogram of echo durations shown in Fig. 6. The mean duration is 107 min. A break of a few minutes has been ignored if the re-appearing echo was at the same location after the break and behaved in the same way (e.g. drifted or remained stationary). A long-period echo which could easily be divided into a period of consistent movement and a stationary period was treated as two echoes.

It was possible, with only a few doubtful cases, to classify the $E_s$ echoes as either drifting consistently or remaining substantially stationary; in the period May–July 1959, 12 days of observation showed ninety-one stationary clouds and fifty-two moving clouds.

5.2. *Statistics of moving clouds*

In studying the statistical behaviour of clouds of $E_s$ it is convenient to represent the position of a cloud at any instant by its mean range and mean azimuth. By following the drift of these co-ordinates for each moving cloud, their mean positions, velocities, and directions of movement have been examined. Fig. 7 shows the available data on the fifty-two moving clouds mentioned above; the arrows are velocity vectors placed at the mean cloud position. A series of histograms of the duration, direction of movement, and velocity of the clouds are shown in Figs. 8 and 9. The two polar histograms (Fig. 8) refer to the number of clouds as a function of azimuth and the number weighted with respect to their duration, respectively. A predominant south-westerly movement is clearly shown. The average duration of a cloud is about $2\frac{1}{2}$ hr, and the average velocity about 220 km/hr; these results are similar to those obtained by PETERSON (1957) at Stanford, U.S.A., where,

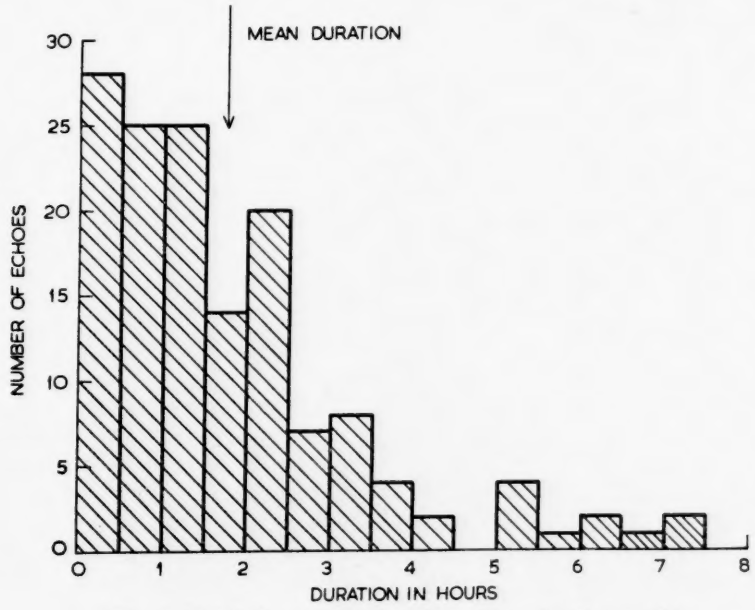


Fig. 6. Distribution of echo durations. Mean duration = $1\frac{3}{4}$ hr.

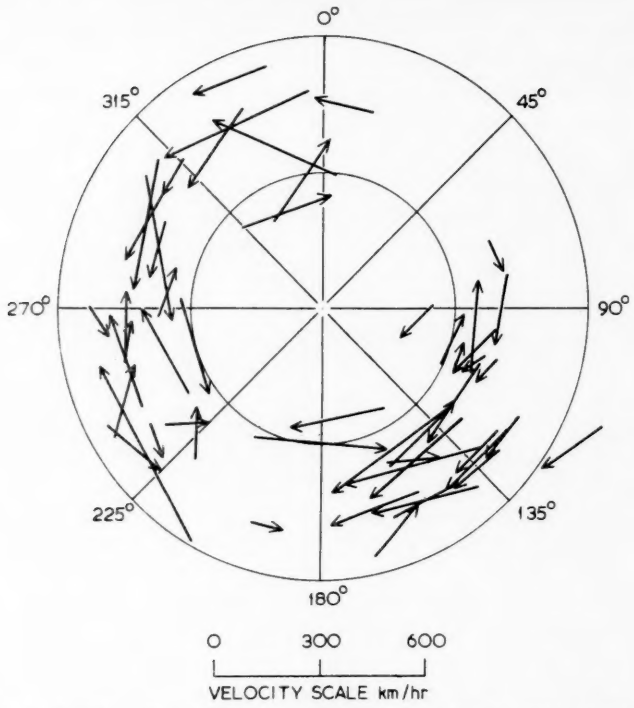


Fig. 7. Position, direction of movement, and velocity of moving clouds.

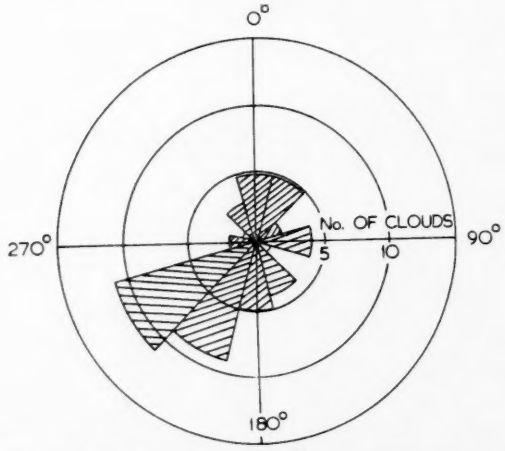


Fig. 8(a). Distribution of directions of cloud movement.

between the years 1952–1955, a predominant westerly movement was shown, the average duration was 3 hr, and the average velocity was 300 km/hr.

A mean south-westerly drift is not in agreement with drift measurements from other sources (BRIGGS and SPENCER, 1954; GREENHOW and NEUFELD, 1956) which show a predominant easterly movement. A few simultaneous drift measurements



VOL. 20 1961

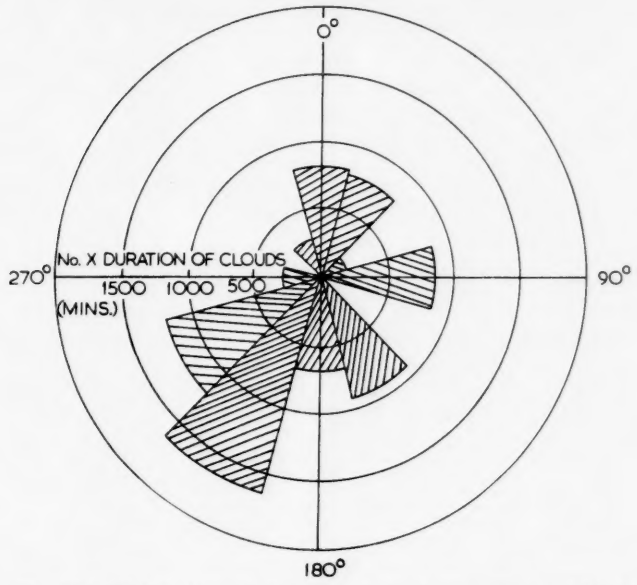


Fig. 8(b). Distribution of direction of cloud movement weighted with respect to cloud duration.

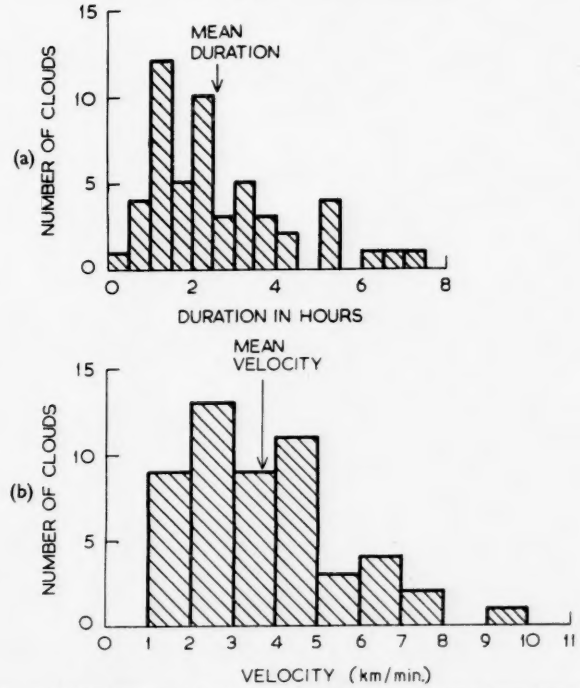


Fig. 9. Distributions of (a) duration and (b) velocity of moving clouds. Mean duration = 156 min. Mean velocity = 221 km/hr = 61 m/sec.

Table 2. Comparison of $E_s$ cloud movement with drift measurements

| Date 1959 | Back-scatter | | | Cambridge drift measurements | | | |
|---|---|---|---|---|---|---|---|
| | Time (hours) | Velocity (m/sec) | Direction (deg.) | Time (hours) | Velocity (m/sec) | Direction (deg.) | Layer |
| | | | | 1120 | 130 | 0 | $E$ |
| 12 May | 1100–1330 | 68 | 250 | 1130 | 180 | 124 | $F$ |
| | | | | 1200 | 120 | 112 | $F$ |
| | | | | 1300 | 111 | 121 | $F$ |
| 13 May | 0740–0900 | 107 | 95 | 1010 | 100 | 72 | $E$ |
| | | | | 0830 | 80 | 90 | |
| 16 June | 0806–1006 | 70 | 340 | 0900 | 90 | 90 | |
| | | | | 0930 | 75 | 54 | |
| | 1052–1300 | 68 | 330 | 1100 | 50 | 40 | $E_s$ |
| | 1115–1230 | 17 | 235 | 1330 | 65 | 180 | $E$ |
| | 1230–1330 | 120 | 235 | | | | |
| | 1312–1340 | 102 | 295 | 1730 | 50 | 220 | $E$ |
| | 1724–1930 | 120 | 245 | 1830 | 50 | 40 | |
| | 2009–2116 | 90 | 230 | 1900 | 50 | 50 | |
| 17 June | 0910–1010 | 112 | 170 | 0930 | 80 | 90 | $E$ |
| | | | | 1000 | 50 | 220 | |
| | 1654–2018 | 33 | 227 | 1700 | 80 | 120 | $E$ |
| 18 June | 2152–2252 | 68 | 225 | 2130 | 80 | 180 | $F$ |
| | | | | 2230 | 45 | 225 | $F$ |
| 19 June | 0533–0903 | 67 | 215 | 0900 | 40 | 90 | $E$ |
| | 0621–0954 | 42 | 225 | 1030 | 45 | 315 | $E$ |
| | 0732–1000 | 47 | 283 | 1330 | 75 | 55 | $E$ |
| | 1340–1510 | 50 | 195 | 1500 | 160 | 0 | $E$ |
| | 1520–1730 | 58 | 60 | 1530 | 65 | 55 | $E$ |
| | 1730–1912 | 83 | 255 | 1630 | 50 | 140 | $E$ |
| | 1821–2330 | 58 | 205 | 1700 | 70 | 30 | $E$ |
| | | | | 1730 | 100 | 50 | $E$ |
| | | | | 1800 | 130 | 90 | $E$ |
| 14 July | 0500–1015 | 40 | 195 | 0500 | 55 | 55 | $F$ |
| | | | | 0600 | 45 | 135 | $F$ |
| | | | | 0625 | 90 | 35 | $F$ |
| | | | | 0900 | 35 | 215 | $E_s$ |
| 15 July | 0630–0930 | 37 | 22 | 0930 | 30 | 135 | $E$ |
| | | | | 1000 | 40 | 160 | $E$ |
| | 0751–1151 | 73 | 163 | 1100 | 35 | 215 | $E$ |
| | | | | 1200 | 65 | 180 | $E$ |
| 16 July | 1728–1830 | 48 | 225 | 1730 | 55 | 30 | $F$ |
| | | | | 1800 | 85 | 230 | $F$ |
| | 1840–2200 | 38 | 83 | 1830 | 55 | 240 | $F$ |
| | | | | 1900 | 65 | 125 | $F$ |
| | | | | 1930 | 90 | 270 | $F$ |
| | | | | 2000 | 60 | 225 | $F$ |
| | | | | 2030 | 80 | 270 | $F$ |
| | | | | 2100 | 35 | 215 | $E_s$ |
| | | | | 2130 | 60 | 230 | $E_s$ |



VOL. 20 1961

for the $E$- and $F$-layers, determined by the three spaced receivers method (BRIGGS and SPENCER, 1954), were kindly provided by workers from the Cavendish Laboratory, Cambridge, and in Table 2 the velocities and directions of drift of $E_s$ clouds are compared with their measurements.

No consistent correlation is evident, even when the drift measurements refer to $E_s$ echoes. It would thus appear that drifts of $E_s$ clouds do not correspond with material drifts or winds, and this view is supported by the fact that simultaneously occurring $E_s$ clouds a few hundred kilometres apart can drift in quite different directions. It is interesting to note that in studies of night-time $E_s$ at Brisbane (THOMAS and BURKE, 1956), a correlation was found between much smaller clouds than those considered here and drift measurements. Thus the possibility is suggested of the motion of a fine structure which correlates with drifts deduced from fading records, whilst the main large cloud moves independently.

## 6. PROPAGATION ALONG THE INVERNESS–SLOUGH PATH

The times of reception and sometimes the amplitude of the pulse received from a transmitter at Inverness on 17 Mc/s were noted. A clear division of the periods

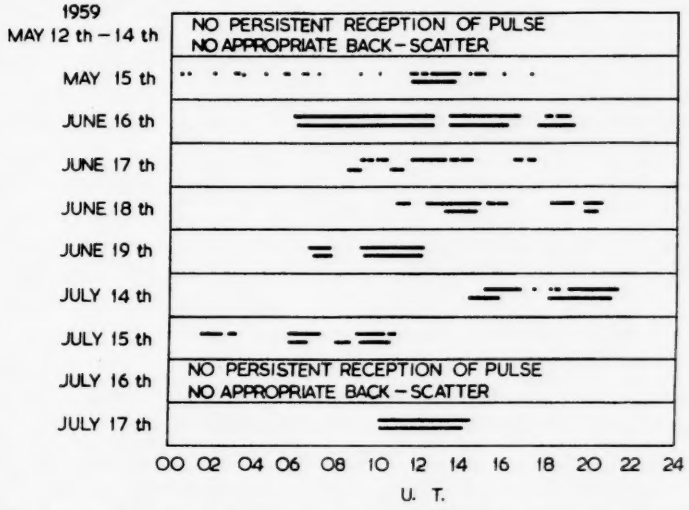


Fig. 10. Correlation between reception of pulse from Inverness (upper trace) and presence of back-scatter (lower trace).

of reception could be made into short-term (a few seconds to a minute) and long-term reception. The former were attributed to propagation via meteor ionization, and the latter to propagation via $E_s$ ionization. At most of the times when a persistent signal was being received, a back-scatter echo also indicated $E_s$ ionization at the mid-point of the path. The correlation between the two sets of information is illustrated in Fig. 10, where the upper trace for each day represents the periods during which a signal was received from Inverness, and the lower trace shows the periods during which back-scatter echoes were received from 700–750 km range

at a centre azimuth within about 20° of 340°, the azimuth of the Inverness transmitter. On two occasions a centre azimuth of about 280° was allowed in view of an exceptionally large azimuthal extent to the echo.

At the times when the signal from Inverness was just appearing or disappearing, it was usually possible to say whether the range or the azimuth change of the $E_s$ cloud observed by back-scatter was the responsible parameter. When it was the range, agreement was within 50 km; when it was the azimuth, the deviation of the centre azimuth from 340° gave a measure of the angular spread of the cloud, which, on the average, was about $\pm 20°$. The corresponding cloud diameter would be about 270 km, a rather larger figure than that deduced from the range measurements.

For the observations in May 1959, a vertical incidence sounder was placed within 1 km of the mid-point of the Slough–Inverness path. The first Fresnel zone in the ionosphere immediately above the sounder had a diameter of about 4 km for vertical incidence reflection at 5 Mc/s, and the corresponding ellipse for the oblique path at 17 Mc/s had a major axis (along the path) of about 12 km and a minor axis of 3·5 km. The sounder was thus sufficiently close to the mid-point of the path for a study of the ionosphere to be made at the expected place of reflection for the oblique transmissions.

Fig. 11 shows the extraordinary ray critical frequency at the mid-point as a function of time, with horizontal lines representing the times of pulse reception at Slough from Inverness; the relative amplitude of the pulse is indicated at intervals by vertical lines. It is clear that for those periods during which $f_xE_s$ was much greater than 5 Mc/s, a strong pulse was received from Inverness, and at other times of day, when $f_xE_s$ was below about 3 Mc/s, the pulse was only received for very short periods via meteoric ionization. A horizontal line is drawn across the critical frequency curve at 5·5 Mc/s. This frequency is calculated, on the assumption that the main propagation between horizontal dipoles at Inverness and Slough is by the extraordinary wave, as the one which the extraordinary ray critical frequency $f_xE_s$ at the mid-point of the path must exceed for propagation to be possible. The calculation assumes the secant law relating the oblique m.u.f. to the critical frequency to be valid for the ordinary ray. The extraordinary ray critical frequency and m.u.f. are both taken at $\frac{1}{2}f_H$ (0·65 Mc/s) greater than their ordinary ray counterparts. It can be seen that the correlation between the times of reception of the pulse and the times for which $f_xE_s$ exceeds 5·5 Mc/s, is reasonably good, but not close enough to allow accurate verification of the secant law; between 1215 and 1236 U.T., for example, the pulse is received at low amplitude from Inverness although the critical frequency is as low as 4·5 Mc/s.

Also shown on Fig. 11 are the smoothed range and centre azimuth plots of back-scatter echoes observed at the same time. On the range plot, the two full curves represent minimum and maximum ranges as shown on the p.p.i., and the dotted curve the maximum range detectable on a more sensitive amplitude display. The far edge of the ground scatter is seen to cover the range of Inverness approximately throughout the duration of the back-scatter echoes. The azimuthal data, however, is more interesting; two drifting echoes are apparent. The Inverness to Slough propagation path is open for back-scatter centre azimuths on the



VOL. 20 1961

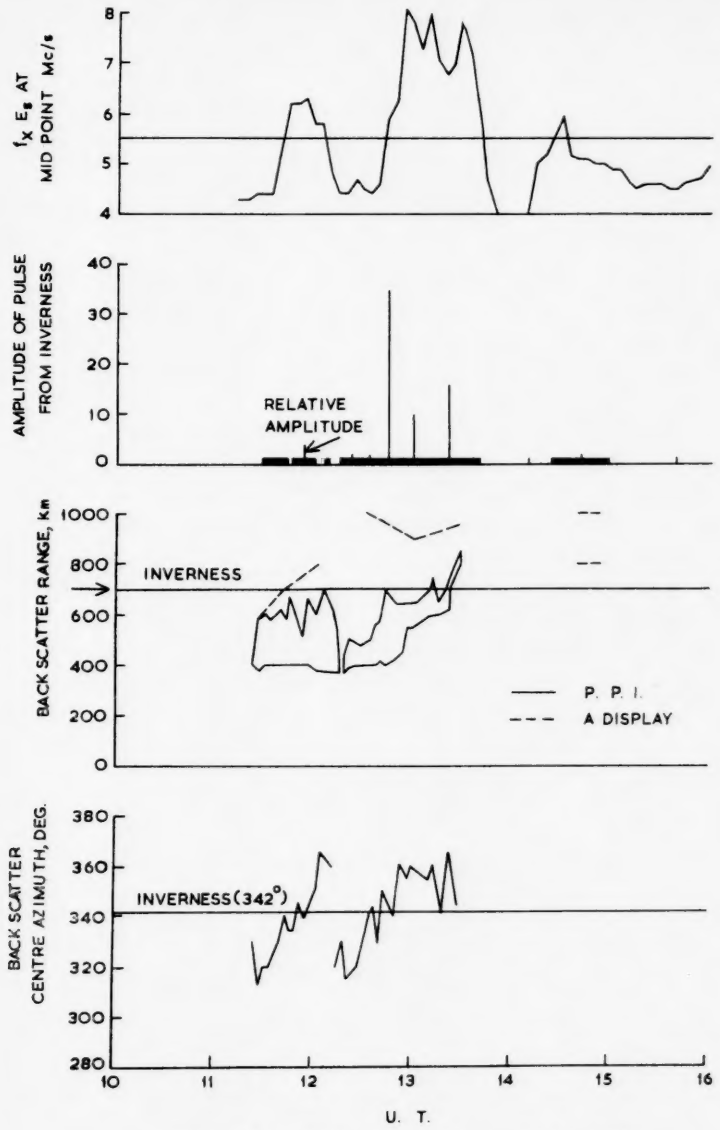


Fig. 11. Sporadic-*E*, 1000–1600 hours, 15 May 1959.

first echo from 320° to 350°, and on the second echo from 320° to 360°. The corresponding cloud width is thus about 250 km, the higher of the figures previously quoted. The last occasion in Fig. 11 on which the pulse was received from Inverness corresponds with the slight increase of vertical-incidence critical frequency, but no back-scatter was seen on the p.p.i. display. The more sensitive A-display showed a very weak back-scatter echo at ranges indicated by the dotted lines. This is

therefore an example of a cloud just seen at vertical incidence, giving a low amplitude one-way transmission from Inverness, and a barely detectable back-scatter echo.

At times on 15 May near to those described above, there were also other back-scatter echoes on the display, and in the next Section the available data are examined and compared with vertical incidence measurements.

## 7. The Distribution of $E_s$ Ionization on Particular Occasions

An attempt has been made in this investigation to augment statistical studies of $E_s$ by a detailed analysis of individual $E_s$ clouds. On 1 or 2 days during the

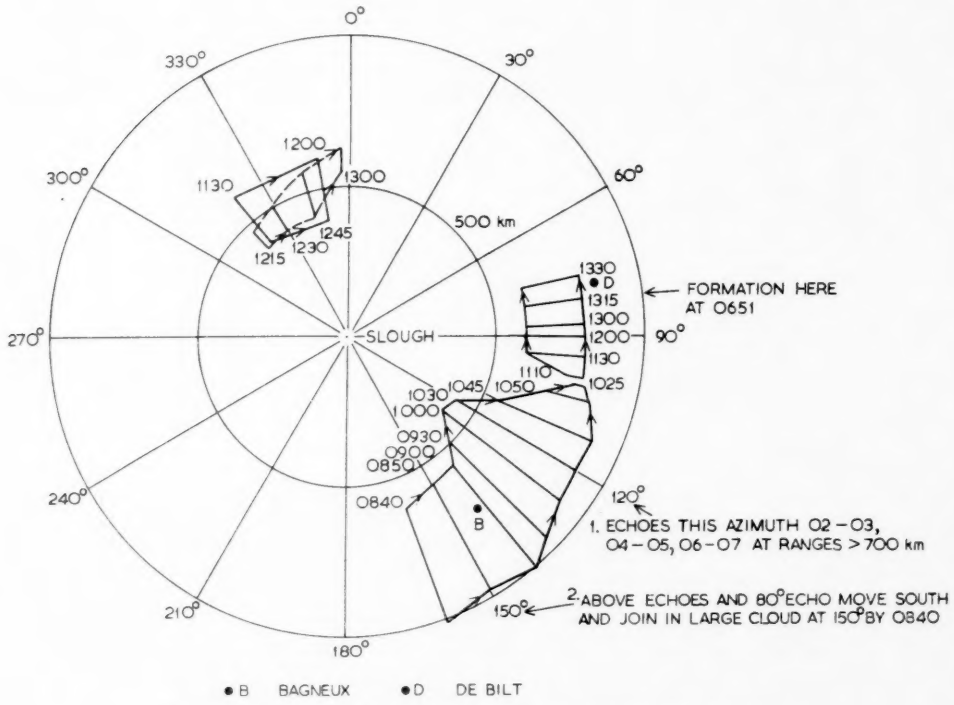


Fig. 12. Movement of $E_s$ echoes on back-scatter p.p.i., 15 May 1959.

experiment, the movement of two or more simultaneous echoes appeared to follow some pattern, and in the following section an attempt is made to present a co-ordinated description of results from one of these days, 15 May.

The essential data obtained on 15 May are summarized in Figs. 12 and 13 which show respectively, the back-scatter data and $f_0E_s$ from various vertical-incidence sounders. During the night (see Fig. 12), relatively weak back-scatter echoes were seen at azimuths near 120° between the hours 0200–0300 and 0400–0500. Slough vertical-incidence records showed $fE_s$ greater than 5 Mc/s during



VOL. 20 1961

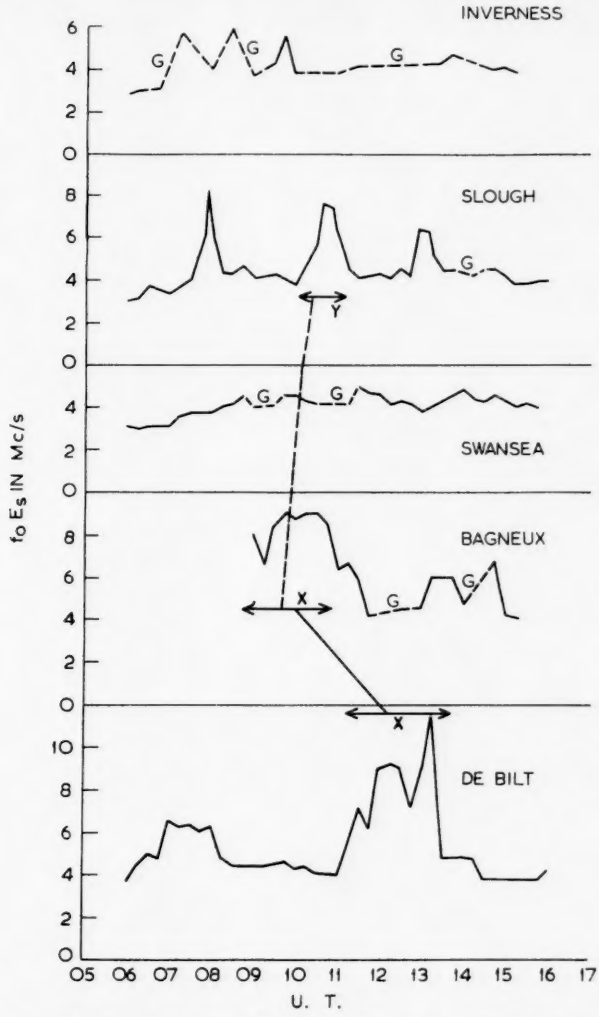


Fig. 13. Vertical-incidence soundings $f_0E_s$ on 15 May 1959.
*XX* $E_s$ occurrence at Bagneux and De Bilt attributable to drifting $E_s$ cloud shown to the south-east in Fig. 12.
*XY* $E_s$ peak at Slough following inward trend of back-scatter range.

the former period. From about 0600 hours, weak echoes moved into the back-scatter field of view at an azimuth of 120°, and at 0651 hours a new echo formed at 80° azimuth. Corresponding to the latter echo, $fE_s$ at De Bilt (82°) rose above 5 Mc/s at 0650 hours, and remained at about 6 Mc/s until 0800 hours, by which time the back-scatter echo had drifted southwards. By 0840 hours, there was one large cloud with a centre azimuth of 160° covering ranges from 600 to 1000 km; this is represented by the line labelled 0840. The movement of this cloud is represented by the subsequent radial lines. At 1100 and 1212 U.T., the cloud

faded out completely but reappeared as though a movement had continued in the interval. The mean speed of cloud movement is 32 m/sec. Correlation of the back-scatter was obtained with vertical incidence results at Bagneux, where $fE_s$ was well above 5 Mc/s from 0900 to 1130 hours, and with De Bilt, where $fE_s$ was well above 5 Mc/s from 1115 to 1330 hours; in fact, the back-scatter results show quite clearly the association of two $fE_s$ peaks displaced in time at the two vertical incidence sounders (Fig. 13). $fE_s$ at Slough showed values above 5 Mc/s between 1030 and 1100 hours, which may be a sequel to the rapid inward trend of the minimum back-scatter range between 0900 and 1000 hours. There are difficulties, however, in using a model of a simple cloud movement without change of form in this instance.

In Section 6, clouds in the northerly azimuths were examined for the period 1124–1330 hours, and it is interesting to note that these also moved in a north-easterly direction, but at a greater velocity of 72 m/sec.

Thus, between the hours of 0840 and 1330, there would seem to be a predominant direction of drift towards the north-east at a velocity between 30 and 70 m/sec. It is not possible to envisage the drift of a simple cloud form during this period, but the data could be interpreted in terms of the drift of a few clouds having linear dimensions of a few hundred kilometres. These clouds occasionally become non-reflecting, and in particular, at 1330 U.T., there appears to be a general decay of the ionization towards the north-east. No back-scatter via $E_s$ ionization was seen on the p.p.i. display for the remainder of the day.

On other days during the experiment some long-term drifts were observed, and these correlated with vertical-incidence soundings at a few stations. It was not possible, however, to fit a simple model to explain the whole of the echo pattern on these days, although parts of the pattern could be so interpreted.

VOL. 20 1961

## 8. Conclusion

The back-scatter technique has been used to examine the location, approximate size, and movement of clouds of sporadic-$E$ ionization for the period May–August 1959 as part of a European co-operative experiment. The data have been correlated with observations at a few vertical incidence sounding stations.

Clouds, of average diameter about 200 km, occurred predominantly in the summer daytime and were seen twice as frequently at latitudes south of Slough as at latitudes to the north. The latitude dependence had previously been noted from vertical incidence soundings, but the actual time for which $fE_s$ exceeded, say, 5 Mc/s on a vertical incidence sounder was much greater than the time for which back-scatter echoes were seen at an appropriate distance from Slough. The difference might be explained in terms of diffuse reflection from a fine structure within a cloud, which, particularly with the smaller clouds, would introduce high attenuation into the back-scatter system. This possibility is supported by the manner in which a typical echo first appeared on the display. Usually a trace occupied a large area in which an echo diffuse in form gradually became more intense; the alternative of a trace of small proportions increasing in size was unusual. The former type of echo could be produced by a change from diffuse to specular reflection as the cloud ionization increased.

During 12 days of observation, 143 clouds were noted. Their average duration was about 2 hr, and a third of them drifted for most of their lifetime (the rest remaining stationary). The predominant drift direction, towards the south-west at a mean speed of 60 m/sec, was not in agreement with drift measurements from other sources, and, no consistent correlation was found between simultaneous $E_s$ cloud movements and $E$- or $F$-layer drift directions on particular occasions. It is therefore suggested that the drift of large $E_s$ clouds does not correspond to the horizontal transport of material particles associated with a true wind, although the fine structure of much smaller clouds or "blobs" might so correspond.

A comparison was made of vertical incidence $E_s$ critical frequencies at the mid-point of the 704 km path from Inverness to Slough, with the times at which pulses at 17 Mc/s were propagated along this path to examine the validity of the secant law relating oblique to vertical incidence propagation frequencies. The correlation was reasonably good, but showed that the secant law was only approximately satisfied for $E_s$, and there was an indication that weak propagation could occur when the critical frequency $f_xE_s$ was about 1 Mc below the expected value of 5·5 Mc/s.

Correlation between the times of reception of the pulse from Inverness, and the presence of back-scatter echoes indicating sufficiently ionized $E_s$ clouds near the mid-point of the path, was good. The start and finish of periods of pulse reception often provided azimuth fixes on the "edge" of an $E_s$ cloud, and using the centre azimuth of the observed back-scatter, an average cloud diameter of about 270 km was deduced.

On 15 May, a predominant drift of echoes in both north-westerly and south-easterly areas was noted; the clouds were moving towards the north-east for about 5 hr at velocities between 30 and 70 m/sec. One of the clouds passed successively over two vertical incidence sounding stations (Bagneux and De Bilt). The $E_s$ critical frequency at these stations was high during the times the cloud, as indicated by back-scatter, was overhead, and in fact, there was a clear correlation between the time-displaced peaks in the curves of $f_0E_s$ against time at the two stations, and the movement of the cloud as shown by back-scatter echoes.

## References


| | | |
|---|---|---|
| Bennington T. W. | 1952 | *Wireless World* **58,** 5. |
| Briggs B. H. and Spencer M. | 1954 | *Rep. Progr. Phys.* **17,** 245. |
| Eckersley T. L. | 1932 | *J. Instn Elect. Engrs* **71,** 405. |
| Gerson N. C. | 1950 | *Nature, Lond.* **166,** 316. |
| Greenhow J. S. and Neufeld E. L. | 1956 | *Phil. Mag.* **1,** 1157. |
| Mitra S. K. | 1954 | Proceedings of the Mixed Commission on the Ionosphere, Brussels, p. 137. |
| Peterson A. M. | 1957 | *Ann. I. G. Y.* **3,** 361. |
| Shearman E. D. R. | 1956 | *Proc. Instn Elect. Engrs* B **103,** 203. |
| Shearman E. D. R. and Harwood J. | 1960 | *J. Atmosph. Terr. Phys.* **18,** 29. |
| Shearman E. D. R. and Martin L. T. J. | 1956 | *Wireless Engr* **33,** 190. |
| Smith E. K. | 1955 | National Bureau of Standards Circular 582. |
| Thomas J. A. and Burke M. J. | 1956 | *Aust. J. Phys.* **9,** 440. |

## APPENDIX

### *The Probability of Occurrence of $E_s$*

DEDUCTIONS have been made in Section 3.2.1 about the probability of finding $E_s$ ionization above a point, from measurements of the occurrence of any back-scatter in the range 700–900 km for 10° azimuthal samples of the data. The relationship between the required and measured probabilities has been deduced as follows.

Let the probability of occurrence of the centre of an $E_s$ cloud in a small area be

$$p(r,\theta)\, r d\theta \,.\, dr.$$

Then the probability $P_1$ that any part of a cloud will be seen above a point $(R,\theta)$ is given by

$$P_1 = p(R,\theta) \,.\, \pi \bar{r}_0^{\,2}$$

where $\bar{r}_0$ = mean cloud radius

Also, the measured probability, $P_2$, that the centre azimuth will be found in $\Delta\theta$, and that any part of the cloud will be found in the range $r_1$ to $r_2$ [where $R = \frac{1}{2}(r_1 + r_2)$] is given by

$$\begin{aligned} P_2 &= p(R,\theta) \,.\, \Delta\theta \int_{r_1-\bar{r}_0}^{r_2+r_0} r\, dr \\ &= p(R,\theta) \,.\, R\Delta\theta\, (r_2 - r_1 + 2\bar{r}_0). \end{aligned}$$

Hence the required probability $P_1$ is related to $P_2$ by

$$P_1 = \frac{\pi \bar{r}_0^{\,2}}{R\Delta\theta(r_2 - r_1 + 2\bar{r}_0)} \,.\, P_2.$$

Putting $\bar{r}_0 = 100$ km

$\Delta\theta = 10°$

$R = 400$ km (back-scatter from 800 km)

$r_2 - r_1 = 100$ km.

Then

$$P_1 \simeq 1{\cdot}5\, P_2.$$

Corrections to this result which might be occasioned by introducing the distribution of cloud radii have been shown to be insignificant.



VOL. 20 1961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 263 to 276. Pergamon Press Ltd. Printed in Northern Ireland

# The *F*-layer at sunrise

H. RISHBETH* and C. S. G. K. SETTY†
Cavendish Laboratory, Cambridge

(*Received* 22 *October* 1960)

**Abstract**—The behaviour at sunrise of the ionospheric *F*-layer has been investigated, using records made at Slough and Cambridge during the years 1953–1958. The rate of increase of electron density at fixed heights, just after *F*-layer sunrise, is greater at sunspot maximum than at sunspot minimum, and greater in winter (December) than in summer (June); the latter effect appears to be associated with the well-known seasonal anomaly in *F2*-layer electron density.

It is suggested that the rate of increase of electron density depends principally on the rate of production of ionization. The seasonal changes might be caused by a change of composition of the upper atmosphere. In the *F*-region electrons are produced by the photoionization of atomic oxygen; the radiation absorbed by molecular nitrogen does not contribute appreciably to the observed ionization, because molecular ions are short-lived in the *F*-region. Consequently a change in the ratio of the concentrations of these gases can alter the effective production rate; it can also affect the loss rate, which depends on the concentration of molecular gases.

In the discussion, vertical transport of ionization is neglected but horizontal movement has to be considered. The calculations of production and loss rates are of limited accuracy, but yield plausible orders of magnitude.

## 1. INTRODUCTION

EARLY DESCRIPTIONS of the sunrise *F*-layer were based on observations of critical frequency, and the equivalent height of reflection of waves at fixed frequencies (APPLETON and NAISMITH, 1935; RATCLIFFE and WHITE, 1934). A seasonal variation was noticed, for the rate of increase of critical frequency was found to be greater, and to begin earlier in relation to ground sunrise, in winter than in summer.

Observations of these kinds, however, refer to heights which change rapidly with time during sunrise, and so are difficult to interpret. More complete information is given by computed height distributions of electron density, or $N(h)$ curves. From a series of $N(h)$ curves can be constructed $N(t)$ curves which show the variation with time of electron density at fixed heights, and which provide information about the rate of production of ionization (RATCLIFFE *et al.*, 1956).

Some observational data about the sunrise *F*-layer at Cambridge and Slough are presented in Section 2 of this paper. Only a first-order explanation of the observed phenomena is attempted, so the observations are not described in great detail. The most important result is that the rate of increase of electron density just after sunrise, $\dot{N}$, is greater in winter than in summer, and at sunspot maximum than at sunspot minimum.

The seasonal anomaly in $\dot{N}$ is connected with the well-known fact that in northern latitudes the noon *F2*-layer electron density is greater in December than in June (e.g. YONEZAWA, 1959). This "winter anomaly" is especially pronounced at Slough (CROOM *et al.*, 1960).

In the analysis contained in this paper, the *F2*-layer is taken to be a "Bradbury layer" (BRADBURY, 1938). At noon, the production peak is supposed to be situated in the *F1*-layer; but at sunrise it is found in the *F2*-layer, owing to the obliquity

* Now at the D.S.I.R. Radio Research Station, Ditton Park, Slough.
† Now at Department of Physics, The Delhi University.

of the solar radiation. Of the assumptions made, the most serious is the neglect of vertical movements due to electromagnetic forces, temperature changes, and plasma diffusion. This is difficult to justify, but the paper shows what conclusions can be drawn if it is valid. These conclusions apply in the first instance to Slough and Cambridge, but may be applicable to other stations in temperate latitudes.

Horizontal movements may be important at sunrise, and are considered in Section 6. Plasma diffusion is neglected, since it can be shown that in some simple cases the observed values of $\dot{N}$ are not significantly affected by its operation (RISHBETH, 1961). With the reservations expressed above, the observations may be discussed in terms of the production and loss processes described in Section 3.

Any satisfactory account of sunrise phenomena in the *F*-region must fulfil certain requirements, which are dealt with in turn in Sections 4 and 5. First, it must suggest an explanation for the seasonal anomalies in $\dot{N}$; and secondly, this explanation must be related to the winter anomaly of noon *F2*-layer electron density. Thirdly, it must be consistent with other data on the upper atmosphere. Fourthly, the deduced rates of production must be consistent with those derived from observations of the noon *F1*-layer. Unfortunately, the quantitative deductions that can be made are of rather limited accuracy.

It is also suggested (Section 5.3) that sunrise observations may be used to obtain very rough values of the loss coefficient.

## 2. OBSERVATIONAL DATA

The observational data on which this paper is based are derived from two sources. The first comprises observations made at Cambridge with a panoramic recorder, over a period of several hours around sunrise, on each of ninety mornings during the years 1954–1958. The second comprises ionograms recorded at Slough from 1953 to 1958, about five hourly records for each of thirty mornings being used, together with Slough $N(h)$ data from the published bulletins. As Slough and Cambridge are only about 90 km apart, the behaviour of the ionosphere at the two places should be very similar.

The records were analysed by the method of THOMAS *et al.* (1957) to yield $N(h)$ curves. From the Cambridge records, these curves were generally computed for intervals of 10 or 15 min, so that fairly detailed $N(t)$ curves could be plotted. Several interesting features emerge from a qualitative study of this data.

On any morning, it is observed that when the Sun's rays strike the *F*-region, the electron density begins to increase and the height of the electron peak to decrease. The $N(t)$ curves vary in shape: sometimes the increase of $N$ is initially slow, later becoming more rapid, and sometimes it is uniform for 2 hr or more. If the finer details are omitted, then some important generalizations can be made from the many examples available:

(a) On most mornings the electron density at a fixed height increases at a fairly constant rate for a period of about 2 hr.

(b) The rate of increase, $\dot{N}(h)$, is at any time greatest at heights near, and somewhat below, the electron peak; and is approximately the same (i.e. to within 30 per cent) at heights within a range of order 50 km or one scale height.

(c) Consequently, it is possible on most mornings to assign a single value of $\dot{N}$, which describes with fair accuracy the rate of increase of $N$ over a range of heights and for a period of at least 1 hr. This will be referred to as the "principal" value of $\dot{N}$.

(d) On some mornings, especially at times of magnetic disturbance, the $N(t)$ curves are rather irregular, and the "principal" value of $\dot{N}$ is less well defined. This may be due to electromagnetic movements, which for the purposes of this paper are neglected.

These conclusions, drawn from Cambridge observations which were made at intervals of a few minutes, show that a good estimate of the "principal" value of $\dot{N}$ can be made from the hourly records customarily obtained by ionospheric observatories, such as Slough.

The use of data from Slough is necessary to overcome a serious inaccuracy of the Cambridge results, which arises from the inability of the equipment to receive echoes at frequencies less than about 2 Mc/s. Because of this limitation, the effect of ionization below the $F$-region is neglected in the computation of $N(h)$ curves, and the resulting errors in the $F$-layer heights may be especially serious near sunrise when the $F2$-layer electron density is least (THOMAS *et al.*, 1957).

Although it would be possible to apply rough corrections to the Cambridge data, the accuracy obtainable would be poor. Better results can be obtained from Slough records, on which $E$-layer echoes are visible soon after sunrise. At times when no $E$-layer trace is recorded, satisfactory corrections to $F$-layer heights may be deduced by using the extraordinary ray (TITHERIDGE, 1959).

Thus the two sets of data are complementary. The Cambridge observations are sufficiently frequent to record the behaviour of the $F$-layer in some detail, but the computed heights are in error because of the neglect of underlying ionization. The Slough data can be corrected for this effect, and so give more reliable measurements of height; the interval between successive records is not too great for the useful determination of $\dot{N}$.

The numerical data are summarized in Table 1. Most of the Slough records analysed are for dates on which Cambridge observations are also available; and although only a few mornings have been used for deriving the average "principal" values $\dot{N}_s$ for Slough, the corresponding Cambridge values $\dot{N}_c$ are representative of larger samples, and so the figures quoted are considered to be typical. $\dot{N}_b$ is the average value of $\dot{N}$ during the 2 hr preceding $F$-layer sunrise: in deriving it, hour-to-hour fluctuations (which may amount to 5 $cm^{-3}$ $sec^{-1}$ at sunspot minimum and 20 $cm^{-3}$ $sec^{-1}$ at sunspot maximum) have been smoothed out. The accuracy of the "principal" values $\dot{N}_s$ is thought to be about 20 per cent.

The most striking feature of the results is that the "principal" values of $\dot{N}$ are greater in winter (December) than in summer (June), and at sunspot maximum (1957/58) than at sunspot minimum (1953–55). Possible interpretations of these observations are discussed in the remainder of the paper, but first some other points in Table 1 should be mentioned.

The column marked $h_s$ gives the range of heights within which the "principal" values of $\dot{N}$ are observed. $\chi_1$ is approximately the value of the Sun's zenith angle at which the increase of electron density begins; soon after this, the $N(t)$ curves

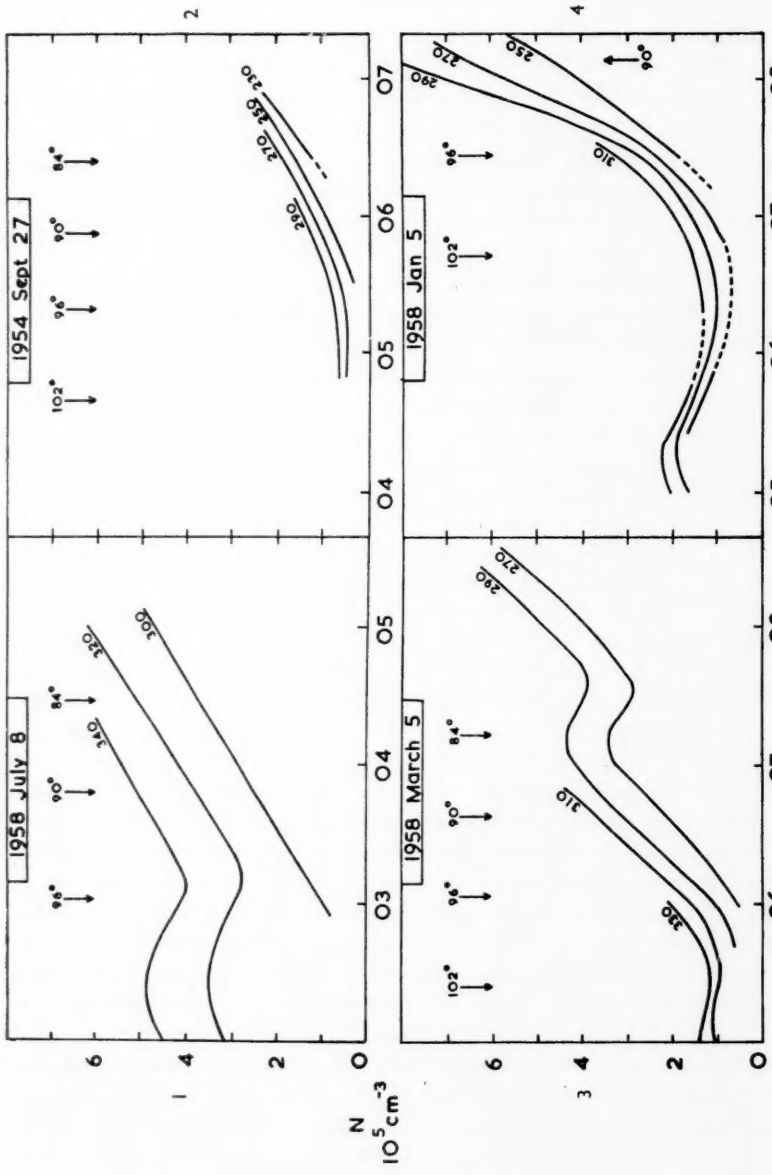


Fig. 1–4. Typical $N(t)$ curves obtained at Cambridge, showing the variation of electron density near sunrise. The heights, in kilometres, are not corrected for underlying ionization (see text). Arrows indicate times at which the Sun's zenith angle attains the values shown.
(1) 8 July 1958; summer at sunspot maximum, showing some pre-sunrise fluctuations.
(2) 27 September 1954; sunspot minimum. (3) 5 March 1958; a disturbed morning.
(4) 5 January 1958; winter at sunspot maximum.



VOL. 20 1961

become approximately linear, at about $\chi = 92°$ in summer and $\chi = 96°$ in winter.

The computed heights for the Cambridge records, in which underlying ionization is ignored, are too great by about 50 km in summer and 25 km in winter at sunspot maximum; the errors are rather smaller at sunspot minimum. No corrections have been made to the heights of the typical $N(t)$ curves shown in Figs. 1–4.

Table 1.

| Epoch and season | $\bar{R}$ | No. of dates | | Rate of increase ($cm^{-3}$ $sec^{-1}$) | | | Slough height (km) | Zenith angle (deg.) |
|---|---|---|---|---|---|---|---|---|
| | | Sl. | Cam. | $\dot{N}_s$ | $\dot{N}_c$ | $\dot{N}_b$ | $h_s$ | $\chi_1$ |
| Sunspot minimum | | | | | | | | |
| Summer 1954 | 0 | 5 | 11 | 15 | 20 | 0 | 220–280 | $92\frac{1}{2}$ |
| Equinox 1954/55 | 0 | 4 | 11 | 14 | 25 | 0 | 200–255 | 97 |
| Winter 1953–55 | 0 | 7 | 7 | 33 | 43 | 4 | 200–230 | $96\frac{1}{2}$ |
| Sunspot maximum | | | | | | | | |
| Summer 1957/58 | 200 | 6 | 13 | 50 | 65 | −5 | 240–340 | $94\frac{1}{2}$ |
| Equinox 1957/58 | 220 | 5 | 5 | 160 | 150 | −5 | 250–305 | 100 |
| Winter 1957/58 | 215 | 4 | 18 | 240 | 200 | 0 | 250–290 | $98\frac{1}{2}$ |

$\bar{R}$ is the mean of the Zürich sunspot numbers for the months in which observations were made.

## 3. Theoretical Considerations

### 3.1. *The continuity equation*

The first step in the discussion is to consider the form taken at sunrise by the equation of continuity

$$\dot{N} = q - \beta N - M \tag{1}$$

in which $\dot{N} = \partial/\partial t\, N(h,t)$; $q$ is the rate of production; and $\beta$ is the loss coefficient. $M$ represents the effects of transport of ionization, and may be written as the divergence of a flux $N\mathbf{v}$.

This equation is linear in $N$, and it is permissible to regard the ionization existing at any time after sunrise as the sum of two components. One (denoted by suffix "$a$") is freshly produced by the incident radiation, being zero before sunrise; the other is the residual night-time ionization (suffix "$b$"). Equation (1) may then be split into two parts:

$$\dot{N}_a = q - \beta N_a - M_a \tag{1a}$$

$$\dot{N}_b = -\beta N_b - M_b \tag{1b}$$

The behaviour of the component "$b$" is in practice irregular, and varies from day to day; but the average values of $\dot{N}_b$ given in Table 1 are not important in comparison to the post-sunrise values of $\dot{N}$. They may without serious error be neglected, and the subsequent discussion is confined to the first component, the suffix "$a$" being discarded.

The interpretation of the values of $\dot{N}$ involves certain assumptions. The loss coefficient is assumed independent of time, though it may vary with season and solar epoch. Temperature changes, which contribute to $M$, are neglected; for it is

doubtful whether large changes can occur within the short period of time with which sunrise observations are concerned. The most serious assumption, however, is that vertical electromagnetic movements do not greatly influence the variation of electron density, and this is difficult to justify, especially since some of the irregularities of actual $N(t)$ curves described in Section 2 might be attributed to such movements. The discussion of horizontal movements is postponed to Section 6.

Plasma diffusion may be important at sunrise (BURKARD, 1957; MARIANI, 1956). This is dealt with in a subsequent paper (RISHBETH, 1961), in which it is concluded that for plausible values of the parameters involved:

(i) The maximum or "principal" value of $\dot{N}$ observed at sunrise is hardly affected by diffusion.

(ii) Diffusion lowers the level of the whole ionized layer, so that the "principal" value of $\dot{N}$ occurs below the level where it would be found in the absence of diffusion. The difference in level is not more than half a scale height.

It is usually sufficient to consider only the simplified equation

$$\dot{N} = q - \beta N. \tag{2}$$

The component of $N$ formed at sunrise (cf. equation 1a) is of course initially zero, so that at first $\dot{N} = q$, and increases as the Sun's rays become less oblique. Later, the $N(t)$ curves become almost linear, partly because the increase of $q$ becomes slower, and partly because the loss term becomes appreciable as $N$ increases. This is discussed in more detail in the subsequent paper (RISHBETH, 1961). The loss term remains of secondary importance throughout the period considered, and so it seems reasonable to suppose that the large changes of sunrise $\dot{N}$ noted in Table 1 must primarily be due to changes in $q$. It is therefore necessary to consider the production process in some detail.

VOL. 20 1961

### 3.2. *Production and loss processes*

The major constituent gases of the $F$-region are probably O and $N_2$; their concentrations will be denoted by $n_1$ and $n_2$ respectively. Up to a certain level $h_d$, which might be in the lower $F$-region, these gases are mixed, so that their scale heights $H_1$ and $H_2$ are equal. Above this level diffusive separation occurs, and the ratio $H_1/H_2$ (which will be denoted by the symbol $k$) exceeds unity; if the gases are in diffusive equilibrium, $k = 1{\cdot}75$, which is the ratio of the molecular weights.

The fundamental production and loss processes may be described in terms of a few important reactions. Solar ultraviolet radiation is absorbed by O and $N_2$: the $N_2^+$ ions are lost rapidly by dissociative recombination, and the $O^+$ ions correspond to the observable electrons. They decay by slower processes involving $N_2$ or $O_2$ molecules, as shown in the following scheme (BATES and MASSEY, 1948):

$$\mathrm{O} + h\nu \rightarrow \mathrm{O}^+ + e \qquad \begin{cases} \mathrm{O}^+ + \mathrm{O}_2 \rightarrow \mathrm{O}_2^+ + \mathrm{O} \\ \mathrm{O}_2^+ + e \rightarrow \mathrm{O}' + \mathrm{O}'' \end{cases} \tag{X}$$

$$\mathrm{O} + h\nu \rightarrow \mathrm{O}^+ + e \qquad \begin{cases} \mathrm{O}^+ + \mathrm{N}_2 \rightarrow \mathrm{NO}^+ + \mathrm{N} \\ \mathrm{NO}^+ + e \rightarrow \mathrm{N}' + \mathrm{O}' \end{cases} \tag{Y}$$

$$\mathrm{N}_2 + h\nu \rightarrow \mathrm{N}_2^+ + e \qquad \mathrm{N}_2^+ + e \rightarrow \mathrm{N}' + \mathrm{N}'' \tag{Z}$$

(Dashes indicate that the atoms may be left in excited states). Although a complete description of the $F$-region must take account of all the above processes, the more restricted situation dealt with here can be adequately discussed in terms of (Y) and (Z) only. It is well known that the type of electron loss formula resulting from a process such as (Y) depends on the concentration of the molecular gas involved, namely $N_2$ (e.g. RATCLIFFE, 1956). In the *F1*-layer and below, a quadratic recombination law applies, the coefficient $\alpha$ being almost independent of height; but in the *F2*-layer the rate of loss is linear in the electron density, with coefficient $\beta = \lambda \, . \, n_2$—where $\lambda$ is the rate coefficient of the atom-transfer reaction. Hence $\beta$ decreases exponentially upward, with scale height $H_2$.

Let $A_1$, $A_2$ be the ionization cross-sections for O and $N_2$, and $\chi$ the solar zenith angle. The production rate $q(h)$ at any height $h$ is $S_h A_1 \, . \, n_1(h)$, in which $S_h$ is the product of $S_\infty$, the flux of ionizing photons incident at the top of the horizontally-stratified atmosphere, and an attenuation factor which depends on the total number of O atoms and $N_2$ molecules in a column of unit cross-section along the path traversed by the radiation. Provided the gradients of scale height are small,

$$q(h) = S_\infty A_1 n_1 \, . \, \exp\left[-n_1 A_1 H_1 \operatorname{Ch}_1\chi - n_2 A_2 H_2 \operatorname{Ch}_2\chi\right]. \tag{3}$$

Here, $\operatorname{Ch}\chi$ is the function defined by CHAPMAN (1931b), and takes different values for the two gases if they are differently distributed; it approximates to $\sec\chi$ when $\chi$ is less than about 85°. $A_1$ and $A_2$ are assumed to be independent of wavelength.

Of the radiation absorbed at any level, a fraction

$$f = \frac{n_1 A_1}{n_1 A_1 + n_2 A_2} \tag{4}$$

is absorbed by O atoms and gives rise to observable ionization: the remainder is expended in process (Z) and thereby lost. The peak value of $q(h)$ occurs where

$$(n_1 A_1 \operatorname{Ch}_1\chi + n_2 A_2 \operatorname{Ch}_2\chi) \, H_1 = 1 + \frac{dH_1}{dh} \tag{5}$$

and is to a first order independent of $dH_1/dh$. If $f_p$ is the value of $f$ at this level, and the difference between $\operatorname{Ch}_1\chi$ and $\operatorname{Ch}_2\chi$ is neglected, the peak value of $q$ is

$$q_p = \frac{f_p S_\infty}{e H_1 \operatorname{Ch}\chi} \, . \, e^g \tag{6}$$

where

$$g = (1 - f_p)(k - 1)/k. \tag{6a}$$

As before, $k = H_1/H_2$; it is assumed independent of height above the level $h_p$ of the production peak. In some circumstances, the correcting factor involving $g$ may amount to 20 per cent or more.

It is convenient to introduce a height coordinate $z$, measured in units of the scale height $H_1$ of the ionizable gas, from the level of peak production for overhead sun. This level will be denoted by the suffix "$o$". If $H_1$ is independent of height, the production function of equation (3) becomes

$$q(z,\chi) = \frac{f_o S_\infty}{e H_1} \exp\left[1 - z - f_o e^{-z} \operatorname{Ch}\chi - \frac{1 - f_o}{k} e^{-kz} \operatorname{Ch}\chi\right]. \tag{7}$$

If $k = 1$ or if $f_0 = 1$, this reduces to the Chapman formula (CHAPMAN, 1931a). The fraction $f_z$ of radiation absorbed by the ionizable gas at any height $z$ varies according to the equation

$$\frac{f_z}{f_0} \cdot \frac{1 - f_0}{1 - f_z} = \exp[(k - 1)z]. \tag{8}$$

The loss coefficient, being proportional to the concentration of $N_2$, varies according to the expression

$$\beta = \beta_0 e^{-kz}. \tag{9}$$

## 4. THE WINTER ANOMALY

### 4.1. *Suggested changes of composition*

It is suggested that the seasonal anomaly in the *F2*-layer is associated with changes of composition of the upper atmosphere, which alter the ratio $n(O)/n(N_2)$, or $n_1/n_2$, in the *F*-region. The results of such a change can be deduced from the equations set out in Section 3.2; the discussion involves the parameter $f$ (equation 4) whose value can be estimated at levels where mixing of O and $N_2$ is complete; i.e. at $h < h_d$. The cross-sections employed are approximate values for wavelengths 500–800 Å; the He II line at 304 Å may also be important (REES and RENSE, 1959).* Relevant data for estimation of $f$ are as follows:

$$\left.\begin{aligned} n_1/n_2 = n(O)/n(N_2) &< \tfrac{3}{4} \quad \text{(WEEKES, private communication)} \\ A_1 &= 1{\cdot}2 \times 10^{-17}\ \text{cm}^2 \quad \text{(BATES and SEATON, 1949)} \\ A_2 &= 2{\cdot}1 \times 10^{-17}\ \text{cm}^2 \quad \text{(EHLER and WEISSLER, 1955)} \end{aligned}\right\} \text{whence } f \leqslant 0{\cdot}3$$

The value of $f$ in the *F*-region could be increased either by an increase in the parameter $k$, or by a lowering of the level $h_d$ at which diffusive separation of O and $N_2$ begins (Section 3.2). There would have to be a seasonal variation of at least 2:1 in $f$ at the level of the *F2*-peak. Suppose that in summer mixing is complete up to the *F2*-peak, so that $k = 1$ and $f = 0{\cdot}3$. Then if in winter the level $h_d$ is lowered to about 150 km, $f$ would increase to 0·9 at the *F2*-peak; and at sunrise, when peak production occurs at this height, the change in $f$ causes a corresponding change in $q_p$ (equation 6) and hence in $N$. But if the value of $f$ is in the first place much less than 0·3 at levels where mixing is complete (as is possible), then it would not be necessary to postulate that the height $h_d$ lies in the *F2*-layer in summer; the required changes in $f$ could occur even if $h_d$ were never greater than about 150 km, provided that $h_d$ varies seasonally by at least one scale height.

A simple illustration of the effects of composition changes may be given as follows. Suppose that with the onset of winter one of these changes takes place at any fixed height in the *F2*-layer:

(A) The concentration $n_1$ of O increases, that of $N_2$ being unchanged;

(B) The concentration $n_2$ of $N_2$ decreases, that of O being unchanged.

* There is some difficulty in regarding this line as the principal source of *F*-region ionization, since its absorption cross-sections in O and $N_2$ (DALGARNO and PARKINSON, 1960) seem to be too small to satisfy equation (5) at the level of the noon *F1*-layer (about 180 km) when recent data on atmospheric density are used.



VOL. 20 1961

Either of these changes would affect the value of $f$, and so give rise to a seasonal variation of sunrise $q$. Their relation to the noon winter anomaly is discussed in Section 4.2. Probably a change intermediate to the types (A) and (B) would agree best with the observed variation of the height of the *F2*-layer, and with the apparent lack of seasonal variation of total atmospheric density implied by satellite observations (KING-HELE and WALKER, 1960).

A change of type (B) decreases the concentration of $N_2$ at any fixed height above $h_d$, and thus reduces the loss coefficient. Now although the loss term in equation (2) is small at sunrise, it is not entirely negligible; and it is found that in some computed examples a decrease of $\beta$ by a factor of 3 may increase $\dot{N}$ by 30 per cent (RISHBETH, 1961). This augments the more important increase due to the change of $q$.

4.2. *Comparison of noon and sunrise seasonal anomalies*

Although both the sunrise and the noon seasonal anomalies could be brought about by the changes described in Section 4.1, they arise in rather different ways. At sunrise, the peak of production lies in the *F2*-layer because of the large solar zenith angle, and for a given $S_\infty$ (which varies with the solar epoch) the peak

Table 2.

| Quantity | Summer $k = 1$ | Winter $k = 1$ | Winter $k = 1{\cdot}5$ |
|---|---|---|---|
| Noon *F2*-peak ($z = 2{\cdot}5$) | | | |
| Ch $\chi$ | 1 | 3·5 | 3·5 |
| $q$ (cm$^{-3}$ sec$^{-1}$) | 246 | 201 | 236 |
| $\beta$ ($10^{-4}$ sec$^{-1}$) | 3·3 | 3·3 | 0·94 |
| $N = q/\beta$ ($10^5$ cm$^{-3}$) | 7·5 | 6·1 | 25 |
| Sunrise (Ch $\chi = 15$) | | | |
| Peak production rate $q_p$ (cm$^{-3}$ sec$^{-1}$) | 80 | | 172 |
| Fraction $f_o$ at $z = 0$ | 0·30 | | 0·30 |
| Fraction $f_p$ at sunrise | 0·30 | | 0·56 |

Assumed values $\begin{cases} S_\infty/(eH_1) = 4000 \text{ cm}^{-3} \text{ sec}^{-1} \text{ (sunspot maximum)} \\ \beta_0 = 0{\cdot}004 \text{ sec}^{-1} \end{cases}$

production at any $\chi$ depends on $f_p$ and $H_1$ (equation 6). The seasonal change in $\dot{N}$ is then attributed to a variation of $f$. But at noon, the peak of production lies in the *F1*-layer; and at the *F2*-peak, the exponential factor in equation (3) is not much less than unity so that $q$ depends primarily on $n_1$. The noon electron density at a fixed height is approximately the equilibrium value $q/\beta$; a change of type (A) described in the previous section increases $q$ by increasing $n_1$ at a fixed height; on the other hand, a change of type (B) reduces $\beta$, and increases $q$ only slightly.

This is illustrated by the example shown in Table 2, which is calculated from the equations obtained in Section 3.2. It is assumed that in summer $k = 1$ and $f = 0{\cdot}3$ at all heights; in winter a change of type (B) occurs, such that $k = 1{\cdot}5$ above the level $z = 0$ while $H_1$ remains unchanged. The noon *F2*-peak is assumed

to be at $z = 2{\cdot}5$; the electron density at this level can be taken as approximately $q/\beta$, notwithstanding the action of plasma diffusion which controls the level of the peak (RISHBETH and BARRON, 1960). The second column is included to show how the noon electron density would decrease from summer to winter, were it not for the change of $k$.

Although this calculation is purely illustrative, it does show how the noon and sunrise effects arise from a change of composition. The parameters are chosen so that the values of noon $N$ and sunrise $q$ are not unlike those observed at Slough, near sunspot maximum.

## 5. DISCUSSION OF NUMERICAL VALUES

So far, the discussion has been mainly concerned with relative values of $\dot{N}$ observed at different seasons, and little has been said about absolute magnitudes of $\dot{N}$. For further consideration, various atmospheric data are required and these are taken from a model atmosphere constructed by Dr. K. WEEKES from densities deduced from rocket and satellite observations. Most of this information has been obtained during the last few years, so that the model represents conditions at sunspot maximum. Two versions of the model are used; in one, complete mixing of O and $N_2$ at all heights is assumed, and in the other diffusive separation of these gases is supposed to occur above 150 km. In the upper $F$-region the gradients of scale height are small ($dH/dh \simeq 0{\cdot}1$ above 220 km), so that the convenient assumption of constant scale height should not be seriously in error.

Table 1 shows that at sunspot maximum the increase of $N$ starts at larger values of $\chi$ than at sunspot minimum. This fact is consistent with an increase of scale height with sunspot number, but more detailed discussion of the quantities $\chi_1$ and $h_s$ requires more complete data about the upper atmosphere than is available at present.

### 5.1. *The flux of ionizing photons*

Attempts to derive numerical conclusions about production rates from the observed values of $\dot{N}$ encounter a difficulty regarding the value of Ch$\chi$ to be used in the equations. At the time when production starts in the $F$-region, $\chi$ is so large that Ch$\chi$ depends very critically upon $\chi$ and $H_1$, and cannot usefully be estimated. By the time $\chi$ has decreased to about 85°, the uncertainty in Ch$\chi$ is greatly reduced, but the loss term is no longer negligible. Perhaps the best compromise is to consider the time at which the $N(t)$ curves become linear, as the loss term is then still small. From values of $\dot{N}$ derived from the calculated $N(t)$ curves described in the following paper, the required value of Ch$\chi$ is estimated as 15, with an uncertainty of at least 25 per cent (which is therefore present in deduced values of the incident photon flux, $S_\infty$): tables (WILKES, 1954) suggest that a larger value might be used.

If the value Ch$\chi = 15$ is adopted, a photon flux of $6 \times 10^{10}$ cm$^{-2}$ sec$^{-1}$ is of the correct order for a sunspot number of 200; this includes the proportion of radiation which is expended in the production of unobserved ionization by process ($Z$) of Section 3.1. With Weekes' data, it gives peak sunrise production rates of 70 cm$^{-3}$ sec$^{-1}$ for the model with complete mixing, and $f = 0{\cdot}3$, and 200 cm$^{-3}$ sec$^{-1}$ for the model in which diffusive separation begins at 150 km. The production peak occurs at 350 km when Ch$\chi = 15$.



VOL. 20 1961

According to the model, the scale height at 350 km is 62 km. This probably varies with the solar cycle, and if corresponding calculations are to be made for sunspot minimum, a smaller scale height is probably appropriate. If a value of 40 km is assumed, the observed summer production rate of 15 $cm^{-3}$ $sec^{-1}$ requires a photon flux of $8 \times 10^9$ $cm^{-2}$ $sec^{-1}$, again assuming $Ch\chi = 15$ and $f = 0{\cdot}3$.

5.2. *Comparison of sunrise and noon production rates*

It is necessary to enquire whether the sunrise production rates $q_p$ are consistent with data on the peak values $q_0$ for overhead sun. Observations of the *F1*-layer at noon at several stations give $q_0$, which is proportional to $S_\infty/H_0$, as a function of monthly mean Zürich sunspot number $R$, such that

$$q_0/\alpha = 500(1 + 0{\cdot}016\ R)\ .\ 10^8\ \text{cm}^{-6} \tag{10}$$

in which $\alpha$ is the *F1*-layer recombination coefficient, of order $10^{-8}$ $cm^3$ $sec^{-1}$. Equation (6) may be applied to the noon and sunrise production peaks, for which suffixes "0" and "$p$" are respectively used, so that approximately

$$q_0/q_p = (Ch\chi)_p\ .\ (f_0 H_p)/(f_p H_0) \tag{11}$$

The comparison thus involves $\alpha$, $(Ch\chi)_p$, and the ratio of scale heights; none of these is accurately known. Further uncertainty arises from possible horizontal movements of ionization, whose effects at sunrise are discussed in Section 6. The most that can be said is that noon and sunrise production rates are consistent, in that equation (11) can be satisfied with plausible values of the various quantities involved: e.g. at sunspot minimum, with $\alpha = \frac{1}{2} \cdot 10^{-8}$ $cm^3$ $sec^{-1}$, $q_0 = 250$ $cm^{-3}$ $sec^{-1}$, $q_p = 15$ $cm^{-3}$ $sec^{-1}$, $(Ch\chi)_p = 15$, and $(f_0 H_p)/(f_p H_0) = 1{\cdot}1$; some of these values might be varied by factors of two.

5.3. *A rough estimate of the loss coefficient*

Although most of the analysis of this paper is based on the idea that the production term in the continuity equation is of primary importance at sunrise, it is possible to make a very rough estimate of the loss coefficient. For this purpose, additional information must be used; namely, the magnitude of $N$ at some time following sunrise, when the loss term has become appreciable. In the specimen calculations shown below, this is obtained from the same data as are used for the compilation of Table 1.

The method depends upon the observation that the "principal" value of $\dot{N}$ remains approximately constant as $\chi$ decreases at sunrise, in spite of the increase of the peak production rate $q_p$ which varies as $(Ch\chi)^{-1}$. It was assumed before, and must be assumed again here, that $Ch\chi = 15$ and $\dot{N} = q$ at the time when the $N(t)$ curves become approximately linear, soon after $N$ first begins to increase. When $\chi$ has decreased to (say) 85° the uncertainty in $Ch\chi$ is much smaller; the tables of Wilkes (1954) give Ch 85° as about 7. It may be deduced that when $\chi = 85°$, the loss coefficient and electron density at the production peak are related by the equation

$$(\beta N)_{85} = \dot{N}\left[\frac{15}{7} - 1\right]. \tag{12}$$

Values of $\beta$ found for the six seasons listed in Table 1 are given in Table 3. Because of the many assumptions made in the calculation, they should not be taken very seriously, but it is interesting that reasonable orders of magnitude are obtained from a wide range of values of $\dot{N}$.

6. *The effect of horizontal movements*

Horizontal movements are generally neglected in analyses of $F$-layer behaviour. This seems reasonable in many situations, but may not be justified at sunrise when

Table 3.

| Season | $\dot{N}_s$ ($cm^{-3}$ $sec^{-1}$) | $N_{85}(F2)$* ($10^5$ $cm^{-3}$) | $\beta$ ($10^{-4}$ $sec^{-1}$) | Approx. height (km) |
|---|---|---|---|---|
| Summer 1954 | 15 | 0·6 | 2·9 | 270 |
| Equinox 1954/55 | 14 | 0·7 | 2·3 | 250 |
| Winter 1953–55 | 33 | 2·0 | 1·9 | 220 |
| Summer 1957/58 | 50 | 2 | 2·9 | 350 |
| Equinox 1957/58 | 160 | 7 | 2·6 | 300 |
| Winter 1957/58 | 240 | 16 | 1·7 | 300 |

* Electron density at production peak, assumed to be about 20 km below $h_m F2$.

there exist large east–west gradients of electron density. If the ionization is drifting westward with velocity $v$, the coordinate $s$ being used to measure distance in this direction, then

$$\dot{N} = q - \beta N - \frac{\partial}{\partial s}(Nv). \tag{13}$$

Suppose that, at all points on the parallel of latitude $\phi$, $N$ varies in the same way with local time. Then the operator $-\partial/\partial s$ may be replaced by $+ (\Omega R_0 \cos\phi)^{-1} \partial/\partial t$, $\Omega$ being the Earth's angular velocity and $R_0$ its radius. Hence

$$\dot{N} = q - \beta N + (N\dot{v} + v\dot{N})(\Omega R_0 \cos\phi)^{-1} \tag{14}$$

so that

$$(1 - \Upsilon)\dot{N} = q - (\beta - \dot{\Upsilon})N \tag{15}$$

where

$$\Upsilon = v \sec\phi/(\Omega R_0). \tag{16}$$

The first-order effect of the drift velocity $v$ is thus to multiply $\dot{N}$ by $(1 - \Upsilon)^{-1}$, $\Upsilon$ being the ratio of $v$ to the peripheral speed of rotation of the Earth, which in latitude 52° is 280 m/sec. There is available much information about the horizontal movement of small-scale irregularities. It is not certain whether these represent actual drifts of the background ionization; but if they do, the correction to $\dot{N}$ may be appreciable. Cambridge data for the period July 1957–September 1958 give an average value of $v$, at sunrise, of about +50 m/sec (B. H. BRIGGS, private communication—see also SHIMAZAKI, 1959), and this increases $\dot{N}$ by about 20 per cent.



VOL. 20 1961

Although no clear seasonal trend is observed, there may be a variation of order 10 per cent in $(1 - \Upsilon)$, which is quite insufficient to account for the winter anomaly in $\dot{N}$, but is in the correct sense to augment the other effects.

The term in $\dot{\Upsilon}$ in equation (15) is probably negligible at Cambridge, where the change in $v$ is only about 20 m/sec per hr. This corresponds to $\dot{\Upsilon} = 2 \times 10^{-5}$ sec$^{-1}$, which is much smaller than $\beta$; and so gives only a small correction to the loss term which is itself of subsidiary importance at sunrise. This conclusion may not be true for lower latitude stations, however, where the gradients of $v$ are much larger (e.g. HIRSH and KNECHT, 1961).

## 7. CONCLUSION

In principle observations of the $F$-layer at sunrise can give information about rates of production of ionization, and hence about the flux of ionizing radiation from the Sun ($S_\infty$), but in practice there are several uncertainties. Consequently the numerical results obtained in Section 5 do not possess great accuracy.

The most striking feature of the observational data for Slough and Cambridge, presented in Table 1, is that the greatest or "principal" rates of increase of electron density after sunrise ($\dot{N}_s$ and $\dot{N}_c$) are greater in winter (December) than in summer (June); and greater at sunspot maximum (1957/58) than at sunspot minimum (1953–55). There might be several explanations of these facts and it is not possible to say with certainty which is correct. In this paper it is suggested that the seasonal changes, which appear to be associated with the seasonal anomaly in noon $F2$-layer electron density, result from changes of the ratio of the concentrations of atomic oxygen and molecular nitrogen in the $F$-region.

The three factors contributing to the variations of $\dot{N}$ are recapitulated in the following list:

I. The peak production rate $q_o$, which is proportional to $S_\infty/H$, increases by a factor of 4·5 as the sunspot number increases from 0 to 220.

II. The fraction $f$ of absorbed radiation which produces observable ionization is altered by a change in the degree of diffusive separation of O and $N_2$, or in the level at which it begins. A change from a condition of complete mixing to one of diffusive separation above 150 km could be accompanied by a change in $f$ from 0·3 (or less) to 0·9.

III. Reduction of $\beta$, which may be another consequence of changes of composition, can increase $\dot{N}$, perhaps by 25 per cent.

Combination of these suggested factors gives a resultant variation of 17:1, sufficient to account for the ratio of the largest and smallest values of $\dot{N}$ in Table 1. But it cannot be claimed that the $F$-region winter anomaly has been "explained", since no cause of seasonal composition changes is known. Nor can any simple explanation be given of the fact that the anomaly is especially pronounced within a certain range of magnetic latitude (CROOM *et al.*, 1960).

It is natural to enquire whether an alternative interpretation of the seasonal anomaly could be given in terms of movements caused by electromagnetic forces. In calculations made for a simple model of the $F$-layer (RISHBETH, 1961), it is found that the "principal" value of $\dot{N}$ is scarcely affected if vertical drift velocities,

of appreciable magnitude, are imposed on the ionized layer: these velocities are assumed to be independent of height but not necessarily independent of time. Thus it might not be a simple matter to explain seasonal variations of $\dot{N}$ solely in terms of vertical transport of ionization, which has been neglected in this paper. In Section 6 it was suggested that $\dot{N}$ might be significantly influenced by horizontal movements, but there is at present no evidence that an alternative interpretation of the seasonal anomaly could be based on this effect.

*Acknowledgements*—The authors wish to express their gratitude to Mr. J. A. RATCLIFFE, F.R.S. for his constant encouragement and many helpful suggestions; and to Dr. K. WEEKES, who read the manuscript and supplied much useful information. This work forms part of the research programme of the Cavendish Laboratory, which is supported financially by the Department of Scientific and Industrial Research. One of us (C. S. G. K. S.) is indebted to the J. N. Tata Endowment, Bombay, for a grant; and to the Government of India and University of Mysore for an overseas scholarship.

## REFERENCES


| | | |
|---|---|---|
| APPLETON E. V. and NAISMITH R. | 1935 | *Proc. Roy. Soc.* A **150**, 685. |
| BATES D. R. and MASSEY H. S. W. | 1948 | *Proc. Roy. Soc.* A **192**, 1. |
| BATES D. R. and SEATON M. J. | 1949 | *Mon. Not. R. Astr. Soc.* **109**, 698. |
| BRADBURY N. E. | 1938 | *Terr. Magn. Atmos. Elect.* **43**, 55. |
| BURKARD O. | 1957 | *Geofis. Pur. Appl.* **37**, 145. |
| CHAPMAN S. | 1931a | *Proc. Phys. Soc. Lond.* **43**, 26. |
| CHAPMAN S. | 1931b | *Proc. Phys. Soc. Lond.* **43**, 483. |
| CROOM S. A., ROBBINS A. R. and THOMAS J.O. | 1960 | *Nature, Lond.* **185**, 902. |
| DALGARNO A. and PARKINSON D. | 1960 | *J. Atmosph. Terr. Phys.* **18**, 335. |
| EHLER A. W. and WEISSLER G. L. | 1955 | *J. Opt. Soc. Amer.* **45**, 1035. |
| HIRSH A. J. and KNECHT R. W. | 1961 | To be submitted for publication. |
| KING-HELE D. G. and WALKER D. M. C. | 1960 | *Nature, Lond.* **185**, 727. |
| MARIANI F. | 1956 | *Ann. Geofis.* **9**, 219. |
| RATCLIFFE J. A. | 1956 | *J. Atmosph. Terr. Phys.* **8**, 260. |
| RATCLIFFE J. A., SCHMERLING E. R., SETTY C. S. G. K. and THOMAS J. O. | 1956 | *Phil. Trans.* A **248**, 621. |
| RATCLIFFE J. A. and WHITE E. L. C. | 1934 | *Proc. Phys. Soc. Lond.* **46**, 107. |
| REES M. H. and RENSE W. A. | 1959 | *J. Geophys. Res.* **64**, 1251. |
| RISHBETH H. | 1961 | *J. Atmosph. Terr. Phys.* **20**, 277. |
| RISHBETH H. and BARRON D. W. | 1960 | *J. Atmosph. Terr. Phys.* **18**, 234. |
| SHIMAZAKI T. | 1959 | *Rep. Ionosph. Space Res. Japan* **13**, 121. |
| THOMAS J. O., HASELGROVE J. and ROBBINS A. R. | 1957 | *Tables of Ionospheric Electron Density*, Series B, No. 1. Cavendish Laboratory, Cambridge. |
| TITHERIDGE J. E. | 1959 | *J. Atmosph. Terr. Phys.* **17**, 110. |
| WILKES M. V. | 1954 | *Proc. Phys. Soc. Lond.* B **67**, 304. |
| YONEZAWA T. | 1959 | *J. Radio Res. Lab.* **6**, 651. |

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 277 to 288. Pergamon Press Ltd. Printed in Northern Ireland

# Diffusion of ionization in the sunrise *F*-layer

H. RISHBETH*
Cavendish Laboratory, Cambridge

(*Received* 22 *October* 1960)

**Abstract**—The variation at sunrise of *F*-layer electron density has been computed for some simple isothermal models of the ionosphere, taking into account the processes of production, loss and plasma diffusion. In some examples, the effects of vertical drifts caused by electromagnetic forces are included, the drift velocities being assumed independent of height but not necessarily independent of time.

It is found that, near the peak of the *F*-layer, the rate of increase of electron density depends primarily upon the production rate, and is not greatly influenced by diffusion or vertical drift. The level at which the peak is found is lowered by the action of diffusion, however, and also depends upon the electromagnetic drift.

## 1. INTRODUCTION

IN TEMPERATE LATITUDES it is observed that when the Sun's rays first strike the *F*-region of the ionosphere, the electron density begins to increase and the height of the electron peak to decrease. The rate of increase of electron density has been used to deduce the rate of production of ionization (RATCLIFFE *et al.*, 1956), but might also depend upon the loss rate and the effects of transport processes. In particular, ambipolar or plasma diffusion plays a large part in the control of the *F2*-layer (e.g. YONEZAWA, 1956) but its effects at sunrise do not seem to have been described hitherto.

Some observational data concerning the sunrise *F*-layer, at Slough and at Cambridge, have been summarized in a previous paper (RISHBETH and SETTY, 1961). In the discussion of these data, possible effects of transport processes were neglected and the seasonal changes attributed to changes of production. The present paper is intended to provide a more complete illustration of the way in which the various processes determine the variation of electron density at sunrise, and to justify the neglect of plasma diffusion in the previous work.

The method adopted for the calculation of electron density, as a function of height and time, avoids the direct integration of the continuity equation. It makes use of the concept—developed by DUNGEY (1956), MARTYN (1956) and DUNCAN (1956)—of distributions of ionization which preserve their shape under the influence of decay processes and plasma diffusion. If distributions of this type are produced in the ionosphere, their subsequent decay and vertical motion can be computed; and at any time a number of such distributions can be superimposed to give a composite layer. This procedure is possible only because, in the *F2*-layer, the equation of continuity is thought to be linear in the electron density.

Some properties of these shape-preserving distributions are summarized in Section 3, after the introduction of the various quantities involved in the continuity equation. The actual calculations are described in Sections 4 and 5; and extended to take account of imposed drifts due to electromagnetic forces, in Section 6. The conclusions drawn from the work are recapitulated in Section 7.

* Now at the D.S.I.R. Radio Research Station, Ditton Park, Slough.

## 2. PRODUCTION, LOSS AND DIFFUSION RATES

The equation of continuity of electron density $N$ is taken to be

$$\dot{N} = q - \beta N - M_D - M_E \tag{1}$$

in which $\dot{N} = \partial N/\partial t$, and $q$ and $\beta$ are the production rate and loss coefficient. $M_D$ and $M_E$ are the terms resulting from vertical movements, due to plasma diffusion and electromagnetic forces. Horizontal transport of ionization is neglected, for the time being.

Throughout this work, the scale height $H$ of the ionizable gas is assumed to be independent of height. The production rate, and the coefficients of loss and of plasma diffusion, can be expressed as simple functions of the coordinates $z$ and $\zeta$, which are measured in units of $H$ from particular zero levels.

The first coordinate, $z$, originates at the level of peak production for overhead sun; and the suffix "o" is used to refer to the level $z = 0$. The other coordinate, $\zeta$, is measured from a level $z = z_s$, which is defined later in terms of the equality of diffusion and loss rates (equation 7), and for which the suffix "$s$" is used. The suffix "$m$" is used to denote the level of the peak of any distribution of ionization.

It is worth noting that in the actual ionosphere the level $z = 0$ is supposed to be at the level of the daytime *F1*-layer, whose altitude is about 180 km; whereas the level $z_s$, or $\zeta = 0$, is two or three scale heights above $z = 0$ and lies near the *F2*-peak.

It is convenient to denote by the symbol $Q$ the function defined by the equation

$$Q(z) \equiv \exp\,(1 - z - e^{-z}) \tag{2}$$

which has a maximum value of unity at $z = 0$. The production function may then be written as

$$q(z,\chi) = q_0 \exp\,(1 - z - e^{-z}\,\mathrm{Ch}\,\chi) \tag{3}$$

$$= \frac{q_0}{\mathrm{Ch}\,\chi}\,Q(z - \log \mathrm{Ch}\,\chi) \tag{4}$$

in which Ch $\chi$ is a complicated function of the solar zenith angle $\chi$, approximating to sec $\chi$ when $\chi < 85°$ (CHAPMAN, 1931a,b). These formulae do not take account of absorption of solar radiation by gases whose vertical distribution differs from that of the ionizable gas (cf. RISHBETH and SETTY, 1961, Section 3.2).

The loss rate is assumed to decrease exponentially with height, so that

$$\beta = \beta_0 e^{-kz}. \tag{5}$$

In all the calculations of this paper, it is assumed that $k = 1$.

The plasma diffusion coefficient $D$ increases with height; it is convenient to define a "diffusion rate" $d(z)$ to be

$$d = d_0 e^z = D/H^2. \tag{6}$$

This function has the dimension of (time)$^{-1}$, the same as that of $\beta$.

The coordinate $\zeta$ is measured from the level at which the functions $\beta$ and $d$ are equal, given by

$$z_s = \tfrac{1}{2} \log\,(\beta_0/d_0) \tag{7}$$



VOL. 20 1961

and so $\beta$ and $d$ may be expressed as functions of $\zeta$; namely, $\beta = \beta_s e^{-\zeta}$ and $d = \beta_s e^{\zeta}$.

## 3. The Behaviour of Chapman-Type Layers

The calculations described in Section 4 depend on the properties of distributions of ionization whose shape is related to the Chapman function $Q(z)$ defined by equation (2). It is convenient to use the term "$Q$-layer" to refer to any layer of ionization whose shape is that of the function $Q(z)$, irrespective of the position of its peak.

The properties of a "$Q$-layer" may be investigated by analysis similar to that applied by Duncan (1956) to the "Chapman alpha" layer whose shape is that of the function $\sqrt{[Q(z)]}$, and which is here described as a "$\sqrt{Q}$-layer".* It may be shown that if a "$Q$-layer" is formed with its peak at the level $\zeta = 0$, it remains stationary and decays without change of shape, with a time constant of $2/(3\beta_s)$. If a vertical drift velocity $w$, independent of height, is imposed on the ionized layer by the action of electromagnetic forces, the stationary position of the peak is at the level

$$\zeta_w = \text{arg sinh } [w/(2H\beta_s)]. \tag{8}$$

A "$Q$-layer" whose peak is situated at a different level will move bodily without change of shape. The equation of motion of the peak of the layer is

$$\frac{d\zeta_m}{dt} = \frac{w}{H} - 2\beta_s \sinh \zeta_m. \tag{9}$$

If $w = 0$, the peak approaches the level $\zeta = 0$, its position at any time $t$ being related to the position $\zeta_m(0)$ at time $t = 0$ by the equation

$$\tanh \tfrac{1}{2}\zeta_m(t) = \tanh \tfrac{1}{2}\zeta_m(0) \,.\, \exp(-2\beta_s t). \tag{10}$$

At any instant, the decay rate of a "$Q$-layer" as a whole is a function of the position of the peak, being

$$b(\zeta_m) = \beta_s [e^{-\zeta_m} + \tfrac{1}{2}e^{\zeta_m}]. \tag{11}$$

This behaviour resembles that of a "$\sqrt{Q}$-layer" which remains stationary at the level where $\beta = \frac{1}{4}d$, and for which the decay rate is simply the value of $\beta$ at the peak; if it is displaced, its motion obeys an equation like (9) (Duncan, 1956).

The second term on the right-hand side of (11) calls for comment, and a brief digression is necessary, in order to point out differences between the layers of shapes $Q$ and $\sqrt{Q}$. The latter is stable in form, under the conditions assumed (Dungey, 1956; Martyn, 1956); and its asymptotic form at great heights, $N \propto e^{-\frac{1}{2}z}$, represents a situation in which the plasma is in diffusive equilibrium under the influence of gravity—provided that the mean mass of the positive ions is the same as that of the neutral atmospheric constituents. But the asymptotic

* The term "Chapman alpha" is often used to describe an ionized layer which is formed by the absorption of monochromatic solar radiation in an isothermal atmosphere, the rate of recombination being of the form $\alpha N^2$. It happens that layers of this shape are important in the quite different context discussed here—viz, plasma diffusion in the $F$-region—and so the term "$\sqrt{Q}$-layer" is used in preference to "Chapman alpha".

form of a "$Q$-layer" is $N \propto e^{-z}$; mathematically, this represents a different solution of the diffusion equation, in which plasma diffuses upwards and is lost to a sink at $z = +\infty$. It seems likely that the "$Q$-layer" shape is unstable, so that if the layer becomes slightly distorted it changes into a form which is in diffusive equilibrium. The additional term in the decay rate $b(\zeta_m)$ may be associated with this instability.

In the analysis of Section 4 changes of this sort are neglected, and the "$Q$-layer" assumed to be stable. This may be justified in that the total period to which the calculations apply, about 2 hr, is probably less than the time $\beta_s^{-1}$ in which changes of shape might become important. Moreover, recent calculations by GLIDDON and KENDALL (private communication) suggest that no great difference is made to electron densities just after sunrise if the boundary condition $N \propto e^{-\frac{1}{2}z}$ as $z \to \infty$ is imposed on the layer, instead of the condition $N \propto e^{-z}$ used here, so that the use of the latter should not introduce any serious error.

## 4. THE SUPERIMPOSITION OF "$Q$-LAYERS"

With these preliminaries, the electron density at times near sunrise may be expressed as a sum of infinitesimal "$Q$-layers". During an interval $\delta t$, a flash of solar radiation is incident at zenith angle $\chi$ and produces an infinitesimal distribution of ionization $\delta N(z) = q(z,\chi) \,.\, \delta t$. This is a "$Q$-layer" with its peak at the level $z = \log \mathrm{Ch}\, \chi$ (cf. equations 3 and 4), and its subsequent behaviour may be traced by using the coordinate $\zeta$ instead of $z$.

The complete electron distribution $N(\zeta,t)$ at any time $t$ is found by superimposing the distributions $\delta N$ formed at all earlier times $t'$, with allowance for their subsequent motion and decay. Residual ionization existing before sunrise is entirely neglected, this being permissible because the equations are linear in $N$. Hence

$$N(\zeta,t) = \int_{t'=t_0}^{t'=t} dN(\zeta,t,t') \tag{12}$$

where

$$dN(\zeta,t,t') = \frac{q_0\, dt'}{\mathrm{Ch}\, \chi(t')} \,.\, Q[\zeta - \zeta_m(t',t)] \,.\, \exp\left\{-\int_{t'}^{t} b[\zeta_m(t',t'')]\, dt''\right\}. \tag{13}$$

In these equations, $\zeta_m(t',t)$ is the level at time $t$ of the peak of the elementary "$Q$-layer" formed at the earlier time $t'$, and may be obtained by integration of equation (9), or—if there is no electromagnetic drift $w$—from the solution given by (10). The limit $t_0$ is the earliest time at which production of ionization is appreciable. The exponential factor in (13) represents the subsequent decay of the elementary "$Q$-layer" produced at time $t'$, $t''$ being a variable of integration; it may be evaluated with the aid of equation (11), once $\zeta_m(t',t'')$ has been computed from (9).

If diffusion is entirely neglected, and there is no electromagnetic drift, the formulae become much simpler. Since the variable $\zeta$ is not applicable in this case, height is measured in terms of $z$. The elementary "$Q$-layers" produced by the incident radiation do not maintain their shape; instead, the ionization produced



VOL. 20 1961

at any height $z$ merely decays with the loss coefficient $\beta(z)$, so that

$$N(z,t) = q_0 \int_{t_0}^{t} \frac{Q[z - \log \mathrm{Ch}\, \chi(t')]}{\mathrm{Ch}\, \chi(t')} \cdot \exp\left[-\beta(z)(t - t')\right] \cdot dt'. \qquad (14)$$

Around the time of sunrise $\chi$ decreases almost linearly with time, and so it is convenient to use $\chi$ as an independent variable in place of $t$, and to refer to the calculated electron distributions as $N(z,\chi)$ distributions, the quantity $(dt/d\chi)$ being treated as a constant. For purposes of numerical calculation, the integrals of equations (12), (13) and (14) are replaced by summations: in most cases, sufficient accuracy is obtained by using an interval of 2° in the summations over the variables $\chi'$ and $\chi''$ (which correspond to $t'$ and $t''$), although on occasion an interval of 1° is necessary.

## 5. The Calculated Electron Density Distributions

For the examples chosen, electron densities have been computed at a series of heights, and at 4° intervals in $\chi$ from $\chi = 94°$ to $\chi = 78°$. At the latitude of Cambridge, +52°, the rate of decrease of $\chi$ near sunrise is about $7\frac{1}{2}°$ per hr at the solstices and 9° per hr at the equinoxes: the adopted value of $(dt/d\chi)$, namely–450 sec/deg., corresponds to a rate of 8° per hr. The parameter $x$, which is required for entry into the tables of Ch $\chi$ (Wilkes, 1954), is the quotient of the geocentric distance of the $F$-region (about 6700 km) and the scale height; the chosen value of $x = 125$ corresponds to $H = 54$ km. The parameter $q_0$ is taken as 400 cm$^{-3}$ sec$^{-1}$, appropriate to a low sunspot number. The level $z_s$ depends upon the assumed values of $\beta_0$ and $d_0$.

Five sets of electron densities $N(z,\chi)$ have been computed for zero electromagnetic drift. They may be divided into two groups. The same loss function $\beta(z)$ is used for each set in the first group, but the diffusion rate $d(z)$ is varied so that $z_s$ takes the values 2·5, 2·0 and 1·5 in the three cases. In the second group, $d(z)$ is the same for each set, but $\beta(z)$ is varied so that the three values of $z_s$ are again 2·5, 2·0 and 1·5. The set with $z_s = 2·0$ belongs to both groups.

The set of $N(z,\chi)$ curves which resembles most closely the $N(h,t)$ curves obtained experimentally (Rishbeth and Setty, 1961) is plotted in Fig. 1. It is the set with the largest loss rate $\beta(z)$, and with $z_s = 2·5$. $N(\chi)$ curves shown for $z = 1·5$, 2·5 and 3·5; the dashed lines are those obtained when diffusion is neglected. High in the $F$-region, at $z = 3·5$, diffusion causes a marked decrease of electron density; at $z = 2·5$, it causes a smaller decrease. For $z = 2·0$, the two $N(\chi)$ curves (not shown in the figure) are very similar, and at $z = 1·5$ diffusion increases the electron density. These effects are also apparent in the $N(z)$ curves for four values of $\chi$ which are plotted in Fig. 2, and which show the manner in which the $F$-layer develops at sunrise. The electron density increases most rapidly near the production peak, whose level is shown for each value of $\chi$, and which lies slightly below the electron peak. It is seen that diffusion lowers the level of the electron peak, by about one-third of a scale height, but reduces the peak electron density by only a small amount (10 per cent in this example).

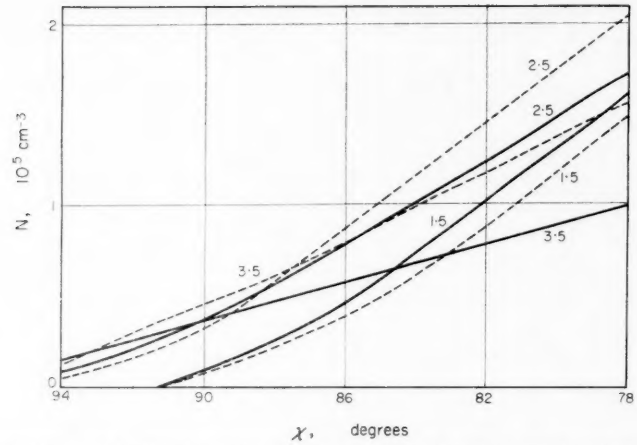


Fig. 1. Variation of electron density $N$ with solar zenith angle $\chi$, calculated for the heights $z = 1{\cdot}5$, $2{\cdot}5$ and $3{\cdot}5$ by the method of Section 4. The curves —— include the effect of plasma diffusion; the curves – – – – – neglect it. Values of parameters are $q_0 = 400$ cm$^{-3}$ sec$^{-1}$; $\beta_s = 1{\cdot}5 \times 10^{-4}$ sec$^{-1}$; $z_s = 2{\cdot}5$.

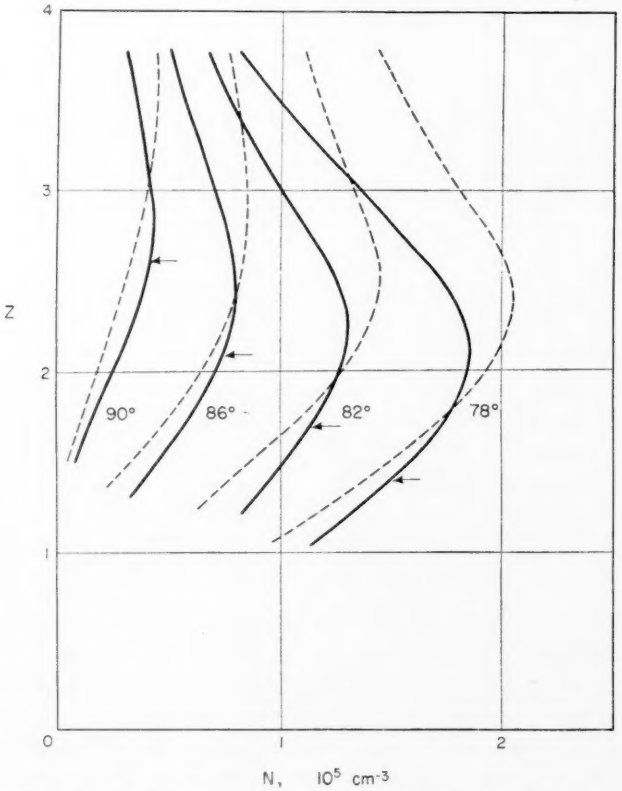


Fig. 2. Electron density distributions $N(z)$ at $\chi = 90°$, 86°, 82° and 78° for the model of Fig. 1. Curves —— include plasma diffusion; curves – – – – neglect it. The arrows indicate the level of the production peak, $z = \log \mathrm{Ch}\, \chi$, at the four values of $\chi$.



VOL. 20 1961

5.1. *Rates of increase of electron density*

The observational data are concerned not only with the electron density $N$, but also with its rate of increase at a fixed height, $\dot{N}$. This quantity is calculated from the finite difference of successive values of $N$. For instance, the value of $\dot{N}(z)$ at $\chi = 84°$ is given with sufficient accuracy by the first-order formula

$$\dot{N}(z,84°) = -\tfrac{1}{4}(d\chi/dt)[N(z,82°) - N(z,86°)]. \qquad (15)$$

Fig. 3 shows graphs of the slopes $\dot{N}(z,84°)$, as functions of $z$, for the examples in the first group described above: the dashed curve refers to the case of no diffusion,

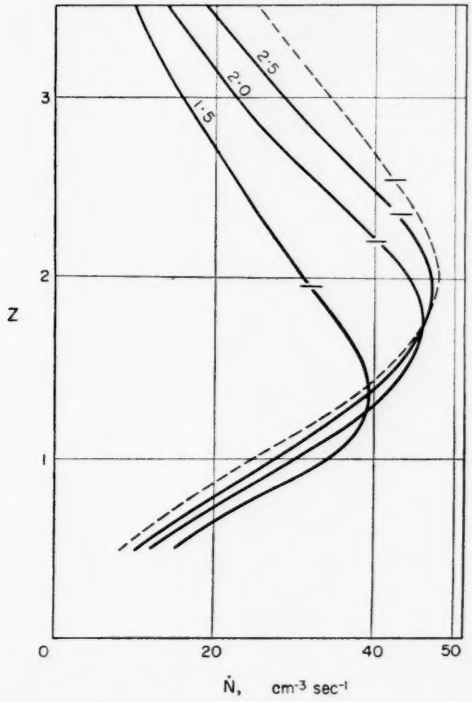


Fig. 3. Rates of increase of electron density at $\chi = 84°$, from the formula of equation (15). For each continuous curve, the same production and loss rates are used ($q_0 = 400$ cm$^{-3}$ sec$^{-1}$, $\beta_0 = 6{\cdot}7 \times 10^{-4}$ sec$^{-1}$), and the diffusion rates are chosen to give $z_s = 1{\cdot}5$, $2{\cdot}0$ and $2{\cdot}5$ in the three cases. For the broken curve, diffusion is neglected ($d_0 = 0$). Horizontal bars denote the level $z_m$ of peak electron density.

in which $\dot{N} = q - \beta N$ only. When diffusion is present, the height at which $\dot{N}$ is greatest is lowered; but the peak value, which will be referred to as $\dot{N}_x$, is only slightly reduced. Although the diffusion rates for the cases $z_s = 2{\cdot}5$ and $z_s = 1{\cdot}5$ differ by a factor of $e^2 = 7{\cdot}4$, the difference in $\dot{N}_x$ is only 15 per cent.

Fig. 4 shows $\dot{N}(z,84°)$ for the second group of examples, which have different loss rates $\beta(z)$ but the same $d(z)$. The dashed curve is the curve of $q(z)$ for $\chi = 84°$, and thus shows how $\dot{N}$ would vary with height if neither loss nor diffusion were present. The peak values $\dot{N}_x$ clearly depend upon the loss rate. A decrease of $\beta$

by a factor of $e$, which reduces $z_s$ from 2·5 to 2·0, increases $\dot{N}_x$ by 40 per cent, but a further decrease of $\beta$ by a factor of $e$ (reducing $z_s$ to 1·5) has a much smaller effect upon $\dot{N}_x$. Although the curves obtained from the equation $\dot{N} = q - \beta N$, which neglects diffusion, are not shown in full, the positions of their peaks are marked with circled points. It is plain that $\dot{N}$ is determined principally by $q$, though modified by $\beta$; it is also plain that diffusion has little effect upon the magnitude of $\dot{N}_x$, but lowers the level at which it occurs by about one-third of a scale height.

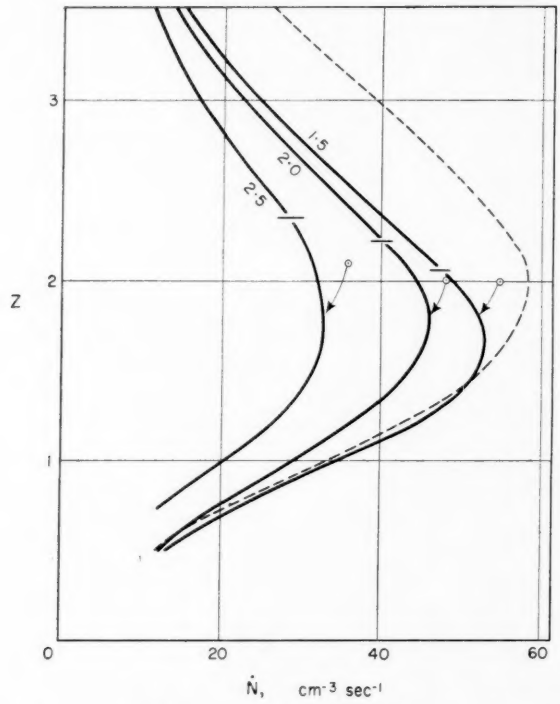


Fig. 4. Rates of increase of electron density at $\chi = 84°$, as functions of height. For each continuous curve, the same production and diffusion rates are used ($q_0 = 400$ cm$^{-3}$ sec$^{-1}$; $d_0 = 1{\cdot}2 \times 10^{-5}$ sec$^{-1}$), and the loss rates are chosen so that $z_s$ is 1·5, 2·0 and 2·5 in the three cases. The circled points show the position of the peaks of the curves $\dot{N}(z)$ obtained with the same values of $\beta$ but with zero diffusion. The dashed curve represents $\dot{N} = q$, with neither loss nor diffusion. Horizontal bars denote the levels $z_m$ of peak electron density.

Curves which are very similar to Figs. 3 and 4 may be drawn for other values of $\chi$; and it is found that at a fixed height the value of $\dot{N}$ varies little with time, this being shown by the linearity of the $N(\chi)$ variation shown in Fig. 1.

As this paper is intended to complement the previous discussion of experimental results, it is interesting to consider these idealized $N(\chi)$ curves in relation to the information obtainable from conventional ionospheric soundings. The levels above the peak of electron density, whose height $z_m$ at $\chi = 84°$ is marked by the horizontal bars on the curves of Fig. 3 and 4, are not accessible to such observations. Now the peak of the $\dot{N}(z)$ curve lies about half a scale height below the electron



VOL. 20 1961

density peak, and so there is an appreciable range of height below $z_m$ in which $\dot{N}$ does not vary rapidly with height. Moreover, it is found that the peak value $\dot{N}_x$ increases only slowly as $\chi$ decreases. This is because the increase of peak $q$, which varies inversely as Ch $\chi$, is balanced by an increase in the loss term.

These features are quite consistent with the observations described in Section 2 of the paper by RISHBETH and SETTY (1961), where it was noted that the observed maximum or "principal" values of $\dot{N}$ remain fairly constant for an hour or more. Indeed, it was suggested in Section 5.3 of the paper that the constancy of $\dot{N}$ could be used to estimate the value of $\beta$.

From the sets of $N(z,\chi)$ distributions described above, and the further sets with electromagnetic drift described in Section 6 below, it is found that the magnitude of the slope $\dot{N}_x$ corresponds to a value of Ch $\chi$ which is about 15. Although this varies somewhat between different examples, and cannot be determined with precision, the value Ch $\chi = 15$ has been assumed for the calculations of Section 5 of the above paper.

### 5.2. *The parameter $x$*

At the beginning of Section 5, the parameter $x$—the ratio of the radius of curvature of the ionosphere to the scale height—was assigned the value 125, corresponding to $H = 54$ km. For this $x$, production of ionization becomes appreciable at $\chi = 97°$. A specimen calculation has been carried out for $x = 50$ ($H = 135$ km); the $N(z,\chi)$ curves are very similar to those described earlier, but differ in that production begins at about $\chi = 101°$. This may be relevant to the seasonal changes noted in Section 2 of the paper by RISHBETH and SETTY (1961).

## 6. The Effect of Electromagnetic Drifts

### 6.1. *Vertical drifts*

The principle of the calculations described in Section 4 is not affected if there exists a vertical drift velocity $w$, independent of height, which is imposed on the ionized layer by the action of electromagnetic forces. In such circumstances, the elementary "$Q$-layers" still maintain their shape, but their motion must be computed by numerical integration of equation (9) instead of from the simple formula (10); the velocity $w$ may be steady, or may vary with time.

Calculations have been made for a model with $z_s = 2{\cdot}5$, the assumed drift velocities being such that, according to equation (8), they produce displacements $\zeta_w$ of the "stationary" level equal to $\pm\frac{1}{2}$ scale height. If $H$ is taken as 54 km (consistent with $x = 125$) the required magnitudes of $w$ are $\pm 8{\cdot}4$ m/sec. The computed $N(z,\chi)$ distributions are very similar to those obtained previously, and illustrated in Figs. 1 and 2, for the case of zero drift. $N(z)$ curves for two values of $\chi$ are shown in Fig. 5, and it is evident that the level $z_m$ of the electron peak is raised by upward drift and lowered by downward drift, and that the peak electron density $N_m$ is virtually unchanged. The peak rates of increase $\dot{N}_x$ are little altered by the drift, details for $\chi = 84°$ being given in Table 1. Corresponding calculations have also been made for the model with $z_s = 2{\cdot}0$.

According to KATO (1956) the vertical drift velocity varies rapidly with time at

sunrise, and the consequences of such a variation have been studied by means of two examples. In one, the rate of change of $w$ is taken as $+8$ m/sec per hr (similar to that suggested by KATO); $w$ is assumed to be zero at $\chi = 90°$. Since $\chi$ decreases by $8°$ per hour, $w$ changes uniformly from $-6$ m/sec at $\chi = 96°$ to $+12$ m/sec at $\chi = 78°$. In the other example, $w$ is reversed in sign but unchanged in magnitude. The rates $\dot{N}_x$ at $\chi = 84°$ are compared in the Table with those computed for cases

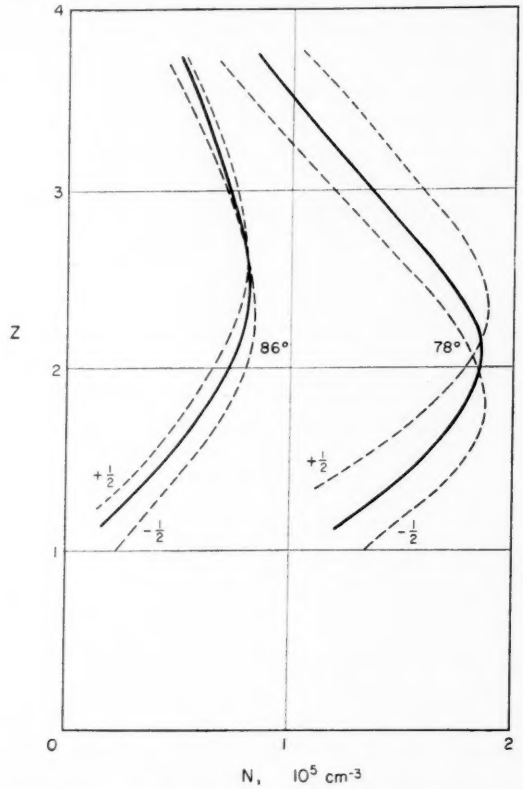


Fig. 5. $N(z)$ distributions at $\chi = 86°$ and $78°$, for the model of Fig. 1 (continuous curves). The broken curves show $N(z)$ distributions obtained when vertical drift velocities, $w$, independent of height and time, are imposed on the layer. The velocities are such as to displace the stationary level by $\pm\frac{1}{2}$ scale height, being $\pm 8{\cdot}4$ m/sec if $H = 54$ km (equation 8).

of steady drift and zero drift. Very similar comparisons may be made at other values of $\chi$. It is concluded that in the simple models studied, vertical drifts have little effect upon the peak values of $N$ and $\dot{N}$, but modify the heights at which they occur.

### 6.2. *Horizontal drifts*

It was shown in Section 6 of the paper by RISHBETH and SETTY (1961) that a horizontal drift $v$ due to electromagnetic forces changes $\dot{N}$ by a factor which depends on the ratio of $v$ to the peripheral speed of the Earth's rotation. This result should



VOL. 20 1961

be unaffected if vertical plasma diffusion is included, as in the analysis of this paper. It should be noted that in horizontal directions plasma diffusion—as distinct from electromagnetic drift—is negligible, even at sunrise.

Table 1. Peak electron densities ($N_m$); peak rates of increase ($\dot{N}_x$); and the heights at which they occur ($z_m$, $z_x$)

| Model | w* (m/sec) | $\zeta_w$ scale hts | $\chi = 84°$ | | $\chi = 82°$ | |
|---|---|---|---|---|---|---|
| | | | $\dot{N}_x$ ($cm^{-3}$/sec) | $z_x$ scale hts | $N_m$ ($10^5$ $cm^{-3}$) | $z_m$ scale hts |
| $z_s = 2{\cdot}0$, $\beta_s = 0{\cdot}91 \times 10^{-4}$ $sec^{-1}$ | | | | | | |
| Steady drift | +11·5 | +1 | 40 | 2·2 | 1·56 | 2·4 |
| | +5·1 | +½ | 44 | 1·9 | 1·61 | 2·2 |
| | 0·0 | 0 | 47 | 1·7 | 1·67 | 2·0 |
| | −5·1 | −½ | 49 | 1·6 | 1·70 | 1·9 |
| | −11·5 | −1 | 49 | 1·4 | 1·67 | 1·8 |
| $z_s = 2{\cdot}5$, $\beta_s = 1{\cdot}5 \times 10^{-4}$ $sec^{-1}$ | | | | | | |
| Steady drift | +8·4 | +½ | 32 | 2·1 | 1·33 | 2·3 |
| | 0·0 | 0 | 34 | 1·8 | 1·32 | 2·2 |
| | −8·4 | −½ | 35 | 1·6 | 1·36 | 2·0 |
| Varying drift: $w = 0$ at $\chi = 90°$ | (Rate of change:) | | | | | |
| | +8 m/sec per hr | | 31 | 2·2 | 1·36 | 2·3 |
| | −8 m/sec per hr | | 37 | 1·6 | 1·38 | 2·1 |

* Assuming $H = 54$ km. Accuracy of $\dot{N}_x$, $N_m$ about 5 per cent; of $z_x$, $z_m$ about 0·1.

## 7. Conclusion

Theoretical curves have been computed to show the variation of electron density at sunrise in an idealized ionosphere. They are similar in form to the $N(t)$ curves commonly observed. The rate of increase of electron density, $\dot{N}$, is greatest at heights just below the electron peak, which are accessible to conventional ionospheric soundings, and remains fairly constant over a period of time. This is similar to the behaviour of the "principal" value of $\dot{N}$ derived from experimental data (Rishbeth and Setty, 1961).

It is found that this greatest value of $\dot{N}$ depends to some extent upon the loss coefficient, but very little upon the assumed value of the diffusion coefficient; it is scarely altered if diffusion is neglected altogether. This provides justification for the neglect of plasma diffusion in the analysis of the above paper, and for the assumption that major changes in $\dot{N}$ must primarily be due to changes in production.

The calculations made in this work are illustrations, based on a very simple model of the atmosphere. But the conclusions reached may apply in the actual $F$-region, even though the height variations of the quantities involved are more complicated than the simple mathematical functions used in this paper. In particular, the loss coefficient may well decrease more rapidly with height than has been

assumed, and this would imply that $k > 1$ in equation (5): DUNGEY (1956) has discussed some distributions of ionization which maintain their shape under such circumstances, and it might be expected that their behaviour resembles that of the distributions appropriate to the case $k = 1$ considered here.

Although vertical drift velocities due to electromagnetic forces may be important in the $F$-region, they seem to have little influence upon the rates of increase of electron density observed at sunrise. This conclusion holds even in some cases in which the drift velocities vary with time, provided that they are at all times independent of height. It thus seems doubtful whether vertical drifts play a major part in determining the behaviour of the $F$-layer at sunrise.

*Acknowledgements*—The author wishes to thank Mr. J. A. RATCLIFFE, F.R.S., and Dr. K. WEEKES for their comments on the draft of this paper; and Mr. J. E. C. GLIDDON and Dr. P. C. KENDALL of Queen Mary College, London, for communicating their recent work (quoted in Section 3 of this paper).

The investigation described forms part of the research work of the Radio Group of the Cavendish Laboratory, supported financially by the Royal Society and the Department of Scientific and Industrial Research.

## REFERENCES


| | | |
|---|---|---|
| CHAPMAN S. | 1931a | *Proc. Phys. Soc. Lond.* **43**, 26. |
| CHAPMAN S. | 1931b | *Proc. Phys. Soc. Lond.* **43**, 483. |
| DUNCAN R. A. | 1956 | *Aust. J. Phys.* **9**, 436. |
| DUNGEY J. W. | 1956 | *J. Atmosph. Terr. Phys.* **9**, 90. |
| KATO S. | 1956 | *J. Geomagn. Geoelect., Kyoto* **8**, 24. |
| MARTYN D. F. | 1956 | *Aust. J. Phys.* **9**, 161. |
| RATCLIFFE J. A., SCHMERLING E. R., SETTY C. S. G. K. and THOMAS J. O. | 1956 | *Phil. Trans.* A **248**, 621. |
| RISHBETH H. and SETTY C. S. G. K. | 1961 | *J. Atmosph. Terr. Phys.* **20**, 263. |
| WILKES M. V. | 1954 | *Proc. Phys. Soc. Lond.* B **67**, 304. |
| YONEZAWA T. | 1956 | *J. Radio Res. Lab.* **3**, 1. |


VOL. 20 1961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 289 to 294. Pergamon Press Ltd. Printed in Northern Ireland

# On the seasonal variation in lunar and solar geomagnetic tides in the geomagnetic equatorial region

K. S. Raja Rao
Meteorological Office, Poona 5, India

(*Received* 4 *November* 1960)

**Abstract**—Following the method outlined by Chapman and Miller (1940), the lunar $L(H)$ and solar $S(H)$ geomagnetic tides at Kodaikanal (geomagnetic latitude $+0{\cdot}6°$) are determined for the three seasons—December solstice, June solstice and equinox, making use of the hourly values of horizontal intensity for the years 1950–1955. The seasonal variation of the amplitude of $L(H)$ at Kodaikanal is compared with similar results for Huancayo (geomagnetic latitude $-0{\cdot}6°$) and Ibadan (geomagnetic latitude $+10°$). It is seen that the amplitude of $L(H)$ at Kodaikanal is larger in December solstice than in June solistice, similar to the variation at Huancayo and Ibadan. It is, therefore, concluded that the ionospheric currents causing geomagnetic tides are not symmetrical with respect to the geomagnetic equator; but the southern hemispheric currents extend into the northern hemisphere up to 10°N geomagnetic latitude in December solstice. It is likely that the region of separation of the northern and southern hemispheric currents is not static, but oscillates northwards and southwards across the geomagnetic equator. The type of seasonal variations in the amplitudes and phases of the lunisolar variations is markedly different from the variations in the higher latitudes. There is no significant seasonal variation in the amplitude of $S(H)$ in the geomagnetic equatorial region.

## Introduction

Chapman and Bartels (1940) have shown that the peculiar monthly cycle of the change of phase of the non-diurnal part of the $L$ field can be adequately explained on the basis of the Dynamo theory. Besides the purely lunar variations, the lunar tide produces other phenomena in which the sun plays a part, giving rise to the lunisolar effects. The chief amongst these lunisolar changes are the ionospheric currents introduced by the lunar tides in the ionosphere owing to the tidal air motion across the geomagnetic field. These reveal themselves in the variations of the geomagnetic field. The variations have a true lunar semi-diurnal component together with other components whose arguments involve both lunar and solar time. The general expression for the lunar diurnal variation may be written as

$$L_n\{2\tau + (n - 2)t + \lambda_n\}$$

where $\tau$ and $t$ are the lunar and solar time respectively and $L_n$ and $\lambda_n$ are the amplitude and phase of the $n$th harmonic. In this expression, the second harmonic $L_2$ of $L$ $(n = 2)$ is purely lunar, not dependent on solar time. But the other components $L_1$, $L_3$, $L_4$ . . . involve both the solar and lunar time and hence they may be called lunisolar components.

## Determination of the Lunar and Lunisolar Tides

Chapman and Miller (1940) have formulated a general statistical method for the analysis of the lunar and solar components of the geomagnetic daily variations and Tschu (1949) has described a practical method of computation. In an earlier paper, Raja Rao and Sivarman (1958) have explained the details of method of computation based on the Chapman–Miller technique of analysis by the fixed age

method. Following this method, the lunar, lunisolar and solar daily variations of the geomagnetic field have been calculated for Kodaikanal. The second lunar components together with their probable errors calculated by Chapman's method (1952) have been represented in the form of harmonic dial in Fig. 1.

The phases are expressed in lunar hours of the occurrence of first maximum. The amplitudes $L_1$, $L_2$, $L_3$, $L_4$ and phases $\lambda_1$, $\lambda_2$, $\lambda_3$ and $\lambda_4$ for all the four harmonics of lunar daily variation are given in Table 1.

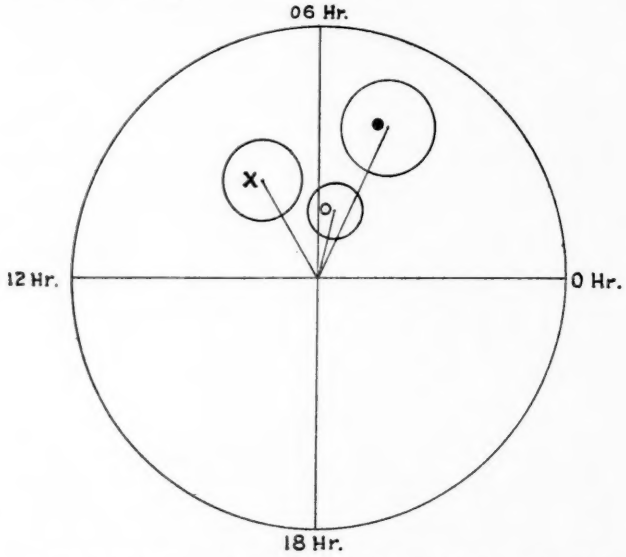


Fig. 1. Harmonic dial for the second components of lunar daily variations, together with their probable error circles. ●, December solstice; × , Equinoxes; ○, June solstice. Radii in $\gamma$ (1 cm = 1$\gamma$).

Table 1. Amplitudes and phases of lunar components at Kodaikanal (amplitudes are expressed in $\gamma$)

| Season | $L_1$ | $\lambda_1$ | $L_2$ | $\lambda_2$ | $L_3$ | $\lambda_3$ | $L_4$ | $\lambda_4$ |
|---|---|---|---|---|---|---|---|---|
| December solstice | 2·20 | 70° 36′ | 3·32 | −39° 04′ | 0·93 | 42° 36′ | 0·30 | 83° 18′ |
| Equinoxes | 1·08 | 31° 45′ | 2·28 | −154° 24′ | 1·50 | 108° 54′ | 0·67 | 59° 46′ |
| June solstice | 0·95 | 12° 48′ | 1·43 | −64° 24′ | 1·18 | 218° 24′ | 0·46 | 119° 02′ |

## Lunar Variations: Discussion of Results

It is seen that at Kodaikanal, as at other places, the second harmonic is the dominant one for the $L$ field. The amplitude of this harmonic is larger in winter (the $D$ solstice) than in summer (the $J$ solstice) by a factor of 2·3. Although Kodaikanal is situated to the north of geomagnetic equator, the seasonal variation of the lunar tide is similar to what is observed at Huancayo which is to the south of the geomagnetic equator.

Onwumechilli and Alexander (1959) have determined the lunar variations of the horizontal intensity of the earth's magnetic field at Ibadan (geomagnetic



VOL. 20 1961

latitude 10·5°N, geographic latitude 7·5°N). They find that the amplitude of $L_2$, the purely lunar component, is greater in winter and smaller in summer. The seasonal variations of amplitudes and phases of $L_2(H)$ at Ibadan, Huancayo and Kodaikanal are shown in Table 2. The position of the three stations and the disposition of the geographic and geomagnetic equators are shown in Fig. 2.

Table 2. Seasonal variation of amplitudes and phases of lunar variation at equatorial stations (amplitudes are expressed in $\gamma$)

| Station | December solstice | | Equinox | | June solstice | |
|---|---|---|---|---|---|---|
| | $L_2$ | $\lambda_2$ | $L_2$ | $\lambda_2$ | $L_2$ | $\lambda_2$ |
| Kodaikanal | 3·32 | −390° 04′ | 2·28 | −154° 24′ | 1·43 | −64° 24′ |
| Huancayo | 8·0 | | 5·9 | | 2·4 | |
| Ibadan | 4·70 | −105° 18′ | 3·82 | −109° 30′ | 3·37 | −109° 30′ |

Large amplitude in $L_2(H)$ in December and small amplitude in June observed at Huancayo, Kodaikanal and Ibadan, indicate the absence of symmetry of the $L$ field about the geomagnetic equator. It is likely that the southern hemispheric currents in the ionosphere causing lunar geomagnetic tides extend into the northern hemisphere at least up to geomagnetic latitude of 10°N in the northern winter. The southern extension of the northern hemispheric current system in June may likewise extend up to 10°S geomagnetic latitude. As we do not have a southern counterpart of Ibadan, it cannot be asserted with definiteness. There is no significant change in phase from winter to summer at Ibadan while at Kodaikanal there is a decrease about 40°. It is, therefore, possible that the region in the ionosphere separating the northern hemispheric and the southern hemispheric currents causing the lunar geomagnetic tides, is not static but oscillates northwards and southwards across the geomagnetic equator.

## Lunisolar Variations

The lunisolar components $L_1$, $L_3$ and $L_4$ are small in comparison with the purely lunar component $L_2$. A comparison of the seasonal variation lunisolar component at Kodaikanal with the seasonal variations at Sitka (Caine, 1957) brings out some interesting facts. In $L_1(H)$ at Kodaikanal, there is a phase decrease from winter to summer by about 50°, while at Sitka also the phase in $L_1(H)$ decreases from winter to summer, but by about 120°. But the amplitude of $L_1(H)$ in Kodaikanal decreases from winter to summer, while the winter amplitudes are smaller at Sitka.

In $L_3(H)$ there is a phase increase by about 176° at Kodaikanal with an increase in amplitude from winter to summer, while at Sitka, the phase decreases by about 180° from winter to summer with an increase in amplitude. In $L_4(H)$ at Kodaikanal there is a increase in amplitude and phase from December to June.

At Ibadan, there is no significant change in the amplitude of $L_1(H)$ from winter to summer; the phase decreases. In $L_3(H)$ and in $L_4(H)$ there is slight increase in amplitude. But there is a decrease in phase by about 34° in $L_3(H)$ and increase by 37° in $L_4(H)$.

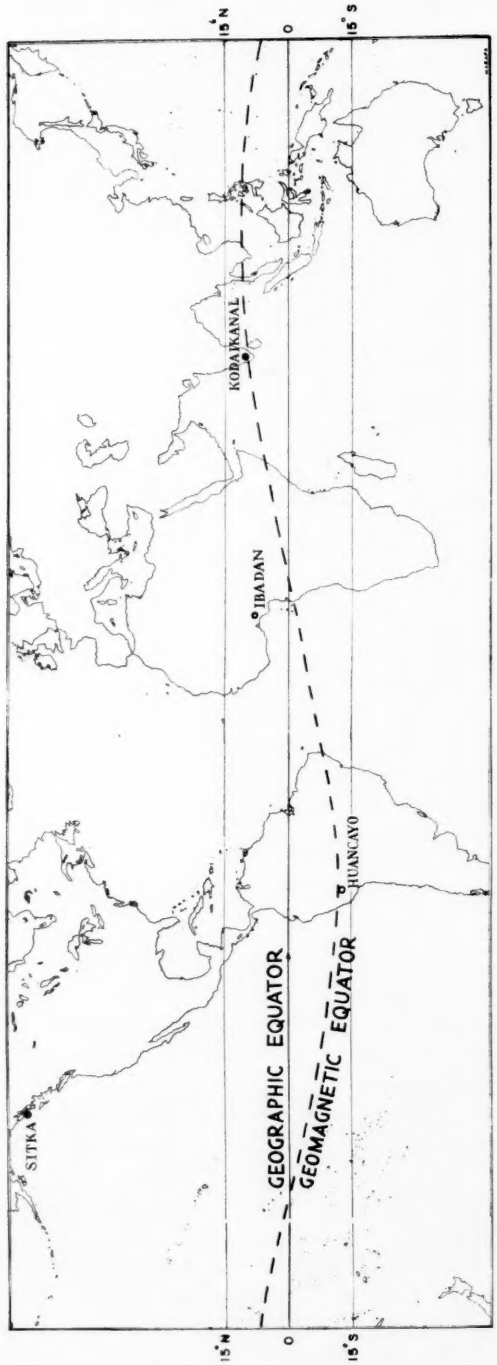


Fig. 2. Location of observatories with respect to geomagnetic and geographic equators.



VOL.
20
1961

From the above considerations, one is led to conclude that the seasonal variation of the $L$ field in the geomagnetic equatorial region is markedly different from similar variation in the higher latitudes. Ibadan appears to lie in the region of transition from the southern hemispheric current system to the northern hemispheric current system, in the ionosphere, responsible for the $L$ geomagnetic tides.

## Solar Diurnal Variations

One of the important features of the Chapman–Miller method of analysis of the $L$ field, is the fact that harmonics of the solar diurnal variation $S(H)$ can also be determined, along with the lunar one. $S(H)$ has been worked out for Kodaikanal up to four harmonics for the winter, equinoxial and summer seasons. The amplitudes ($S$) and phases ($\sigma$) are given in Table 3.

Table 3. Amplitudes and phases of solar daily variation at Kodaikanal during $D$, $E$ and $S$ seasons (amplitudes are expressed in $\gamma$)

| Season | $S_1$ | $\sigma_1$ | $S_2$ | $\sigma_2$ | $S_3$ | $\sigma_3$ | $S_4$ | $\sigma_4$ |
|---|---|---|---|---|---|---|---|---|
| December solstices | 19·1 | −169° 31′ | 14·6 | 234° 34′ | 7·8 | 208° 27′ | 2·2 | 24° 00′ |
| Equinoxes | 24·5 | −73° 47′ | 18·9 | 118° 48′ | 10·0 | 134° 41′ | 4·3 | 113° 00′ |
| June solstices | 22·6 | −75° 00′ | 14·8 | −67° 12′ | 5·6 | 218° 24′ | 4·2 | 63° 18′ |

Table 4.

| Station | Season | | |
|---|---|---|---|
| | December | Equinox | June |
| Huancayo | 5·1 | 8·8 | 16·5 |
| Kodaikanal | 6·0 | 10·7 | 15·8 |
| Ibadan | 8·3 | 12·1 | 10·8 |

As at any other place, $S_1(H)$, the first harmonic is the most dominant one. It has almost the same magnitude in December as in June, both in amplitude and phase, similar to Huancayo and Ibadan.

## Comparison of $S$ and $L$ at the Equatorial Stations

Onwumechilli and Alexander (1959) have taken $S_1(H)$ with $L_1(H)$ and $S_2(H)$ with $L_2(H)$ for purposes of comparison between the $S$ and $L$ fields. A better physical picture can be obtained by restricting the comparison to the predominant harmonic of $S(H)$ with the most predominant harmonic of $L(H)$. The ratios of $S_1(H)$ to $L_2(H)$ are given in Table 4 for the three equatorial stations for the three seasons.

Comparison of $S$ and $L$ at these equatorial stations brings out the following main features. The major difference between $S$ and $L$ is the larger size of $S$. The ratio of the amplitudes $S_1/L_2$ is a function of geomagnetic latitude. As there is no appreciable seasonal variation in $S$, the large seasonal variation in $L$ is reflected

in the increase of the ratio $S_1/L_2$ from December to June. In the higher latitudes, amplitude changes with season in $S$ are much larger than those of $L$ (Caine, 1957).

## Conclusion

The study of the lunar and solar harmonic analyses at Kodaikanal, in comparison with similar observation for other geomagnetic equatorial stations, reveals that the seasonal variation in $L_2(H)$ up to $+10°$ geomagnetic latitude is similar to the variation just to the south of geomagnetic equator. The southern hemispheric currents in the ionosphere extend up to 10°N in northern winter. It is also likely that the region separating the northern and southern hemispheric currents oscillates across the geomagnetic equator. The seasonal variations in the lunisolar components is markedly different in the lower latitudes from what it is in higher latitudes. Comparison of $S_1(H)$ and $L_2(H)$, shows that in the equatorial region amplitude changes with season in $S(H)$ are much smaller than those of $L(H)$ while in higher latitudes the reverse is the case.

*Acknowledgement*—The author is grateful to Mr. R. Challappa for his valuable help in computations.

## References


| | | |
|---|---|---|
| Caine J. C. | 1957 | Report No. AF19 (604)-503, Geophysical Institute, University of Alaska. |
| Chapman S. | 1952 | *Austr. J. Sci. Res.* **5**, 218. |
| Chapman S. and Bartels J. | 1940 | *Geomagnetism.* Oxford University Press. |
| Chapman S. and Miller J. C. P. | 1940 | *Mon. Not. R. Astr. Soc., Geophys. Suppl.* **4**, 649. |
| Onwumechilli C. A. and Alexander N. S. | 1959 | *J. Atmosph. Terr. Phys.* **16**, 106. |
| Raja Rao K. S. and Sivarman K. R. | 1958 | *J. Geophys. Res.* **63**, 727. |
| Tschu K. K. | 1949 | *Austr. J. Sci. Res.* **2**, 1. |




VOL. 20 1961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, pp. 295 to 299. Pergamon Press Ltd. Printed in Northern Ireland

# RESEARCH NOTES

## The effect of multiple atmospheric inversions on tropospheric radio propagation

(*Received* 28 *November* 1960)

Of the various mechanisms put forward in recent years to explain long range tropospheric propagation the most important seem to be scattering from atmospheric turbulence and partial reflection from high level subsidence inversions. In 1952 the writer, after first reviewing earlier experimental work, showed theoretically that a well-developed inversion could cause the phenomenon (Northover, 1952). In a subsequent three-part paper (Northover, 1955), after first reviewing the major theories which had been put forward in the interim, he considered in greater detail the special characteristics of the field under a high level subsidence inversion (Part II). It is now fairly generally agreed that the major mechanisms responsible for the long distance fields in tropospheric propagation are scattering from atmospheric turbulence, and partial reflection from high level subsidence inversions.

Refractometer measurements have shown that several minor high level inversions can co-exist at different levels and that this situation is probably present for most of the time. It therefore seemed to the writer that the extension of the analysis referred to above to take account of the case when several inversions are present, would be of interest. The full analysis (apart from some algebraical detail) is given elsewhere (Northover, 1961); in this note we summarize the results.

The following "addition theorem", which makes it possible to deduce the field under a system of multiple inversions from the standard analysis already given (Northover, 1952, 1955) for a single inversion, was discovered. Here, $R$ is the earth's radius and $\lambda$ is the wavelength of the radio waves used.

"When the inversion layers are at heights large, or moderately large compared with $R^{1/3}\lambda^{2/3}$, and when also they are sufficiently thin to act effectively as dielectric discontinuities then the field under the system is given by the analysis of Northover (1952), with the value of '$k$' therein taken as the sum of the values of $k$ appropriate to each inversion separately."

The conditions of this theorem embrace most situations of importance in practice and when they are satisfied the propagation is governed by the analysis already developed by Northover (1952).

This note is concluded by remarking that there is one special case of the theorem in which the characteristic properties of the propagation can be clearly perceived. It occurs when the layers are close together (i.e. when their widest separation is small compared with the height of the lowest). In this case the writer has shown that the propagation is actually equivalent to that which would be caused by a single inversion at the mean height having a dielectric discontinuity equal to the algebraic sum of the discontinuities of the separate inversions.

In the writer's 1952 paper, it was found that a sharp high level inversion whose dielectric constant difference was about $2{\cdot}58 \times 10^{-5}$ could cause considerable anomalous propagation of metre radio waves. Inversions of this size do not, perhaps, often occur, but under settled weather conditions it is more likely that sharp high level inversions of much smaller

dielectric constant difference would occur much more frequently. If several such small inversions occurred simultaneously (e.g. four such that $\delta_j - \delta_{j+1} = 0{\cdot}6 \times 10^{-5}$ would do), the theorem shows that propagation of the waves well below the inversions would be effectively the same as that under a single large inversion. A mechanism has therefore been found which is capable of explaining the long distance fields without having to invoke the existence of a single strong elevated high level inversion layer.

*Department of Mathematics*
*Carleton University, Ottawa*

F. H. NORTHOVER

REFERENCES

NORTHOVER F. H. 1952 *J. Atmosph. Terr. Phys.* **2,** 106.
NORTHOVER F. H. 1955 *Canad. J. Phys.* **33,** 241, 316,
NORTHOVER F. H. 1961 *J. Res. Nat. Bur. Stand.* D. In press.

---

## Travelling disturbances in the *F*-region over Waltair

(*Received* 30 *November* 1960)

VOL. 20 1961

TRAVELLING disturbances in the *F*-region have been studied extensively by MUNRO (1950, 1953) who has utilized the effect of these disturbances on the *P'–t* records at three widely spaced points to study their horizontal movements. Abnormal patterns in the ionosonde records reported from some low latitude stations such as Singapore (HON YUNG SEN, 1949; OSBORNE, 1951) and Ibadan (SKINNER *et al.*, 1954) were considered to be due to thick layer structure and stratification, but MUNRO and HEISLER (1956) showed that these may be interpreted as due to a change of distribution of ionization due to the passage of a travelling disturbance. In this communication the authors present a description of the general characteristics of travelling disturbances as observed at Waltair on the *P'–t* and *P'–f* records.

Observations of the effects of travelling disturbances on the *P'–t* records were taken at 6·0 Mc/s and on *P'–f* records were taken mainly in the frequency range 5–15 Mc/s almost every day between 0800 and 1800 hours during the period February 1956 to January 1958. From the vast volume of data available it is observed that the travelling disturbances occur more frequently during afternoon hours at this station. The seasonal variation was found to be similar to that observed at a high latitude station by MUNRO, in that the occurrence of these disturbances was more frequent during summer and winter seasons. However, during the winter months the travelling disturbances occur much more frequently as there are no occasions when they could not be observed during daytime.

It is interesting to note that a quick periodic fading is invariably observed on the single reflected echo before it splits and after the split components merge, while a travelling disturbance passes over the observing station. These two components have usually a virtual height separation of 10–300 km, which is quite large compared to the 10–100 km virtual height separation reported by MUNRO. It was noticed by the authors that larger separations are observed generally in winter months. Fig. 1(a) shows a *P'–t* record for which a maximum virtual height separation of 250 km was observed. It may be noticed that this separation is widely variable from time to time. In some other records the separation between the two split components shows cyclic variation with the duration of each

OL.
20
961

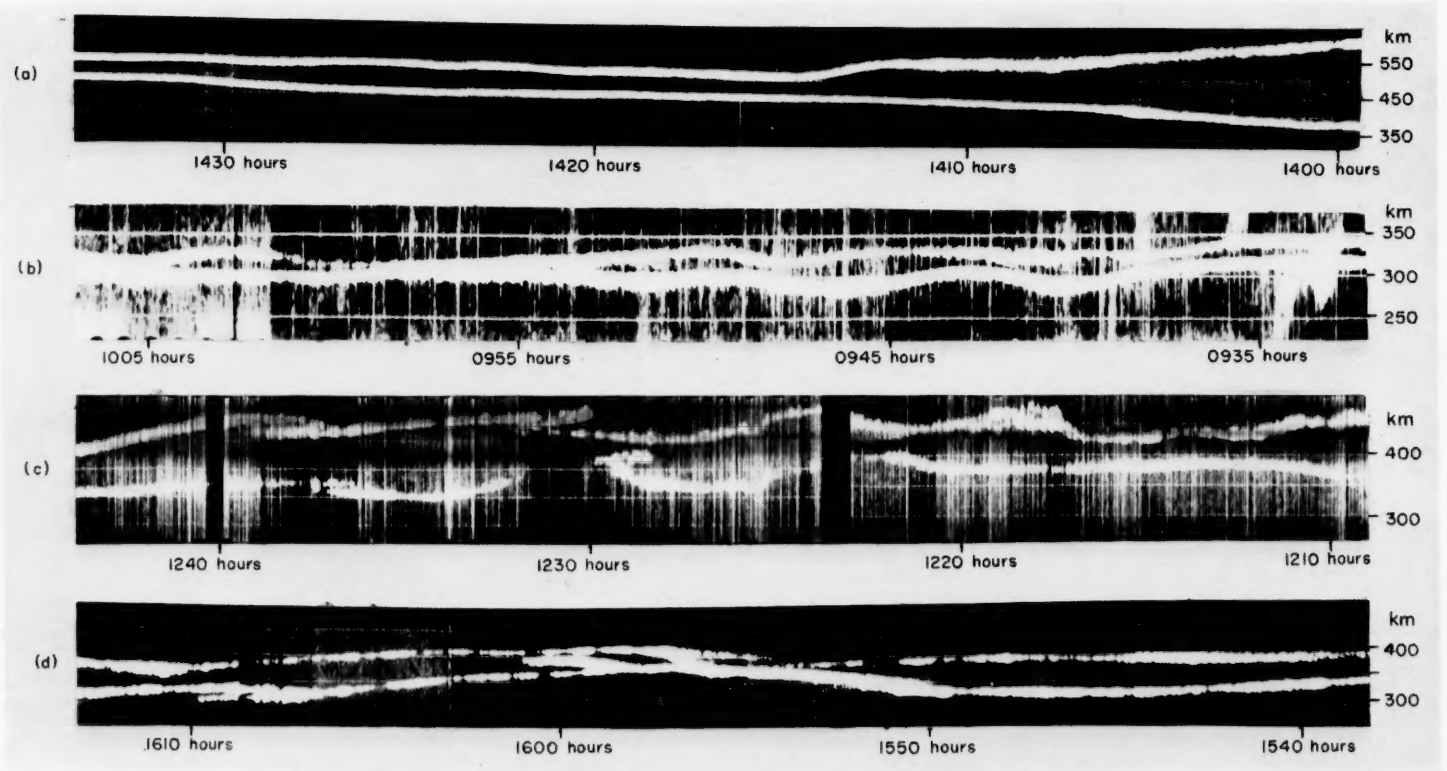


Fig. 1. $P'$–$t$ records showing effects of travelling disturbances. (a) 30 December 1957, 7·0 Mc/s. (b) 21 July 1957, 6·6 Mc/s. (c) 20 June 1957, 6·0 Mc/s. (d) 1 January 1958, 6·0 Mc/s.

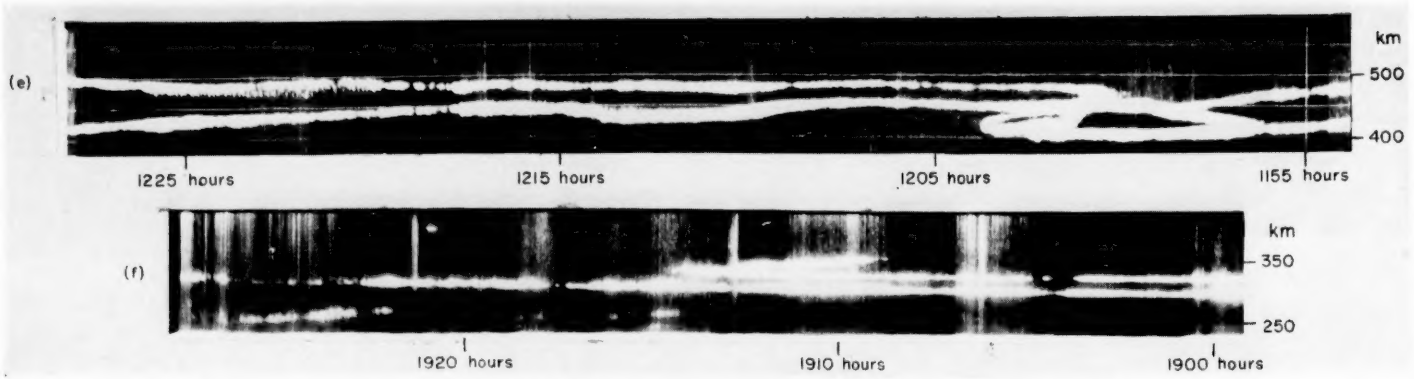


Fig. 1(cont.). $P'$–$t$ records showing effects of travelling disturbances. (e) 28 December 1957, 7·0 Mc/s. (f) 27 January 1957, 2·1 Mc/s.

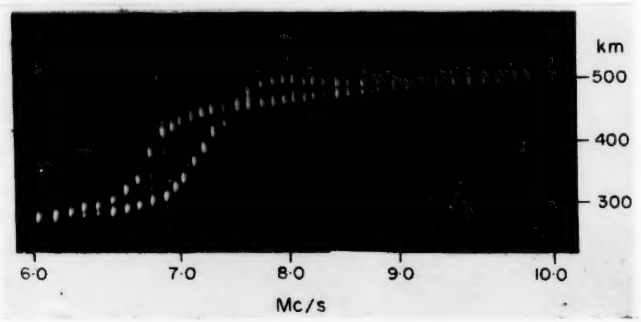


Fig. 2. $P'$–$f$ record, 1450 hours, 30 July 1958.

cycle varying generally from 5 to 15 min. Very often several cycles occur in succession with or without any cross-over before merging.

Fig. 1(b) shows part of a record in which two such cycles occur in succession followed by a cross-over.

$P'$–$f$ curves obtained manually were found to be in general similar to those reported by MUNRO. The $P'$–$f$ record reproduced in Fig. 2 shows the unusual feature of a large group retardation in both the split components at two different frequencies near 7 Mc/s. The change in the virtual height is of the order of 200 km. Polarization tests were carried out on both these split echoes and it is found that the echo corresponding to higher equivalent height at lower frequencies is ordinary whereas at frequencies above 7·5 Mc/s the extraordinary echo occurs with the higher equivalent height. In all the cases where the main echo splits into two it is found invariably that one of them is ordinary and the other extraordinary and hence they are called $o$ and $x$ rays. Short time perturbations of the type reported by MUNRO known as $x$, $y$ and $z$ type complexities were quite often noticed at this station on both the $o$ and $x$ rays but there were occasions when they were observed on either $o$ or $x$ rays alone. The occurrence of $x$ type complexity is rather unusual at this station. Some of the $P'$–$t$ records showing these complexities are shown in Figs. 1(c), (d) and (e). Fig. 1(c) shows some of the $y$ and $z$ complexities occurring on both the echoes within a short interval of time. Fig. 1(d) shows two $y$ type complexities occurring on the same echo. It is interesting to see that the earlier one occurs on the higher virtual height side whereas the later one occurs on the lower virtual height side. An unusual type of complexity which has not so far been reported is reproduced in Fig. 1(e). From the sequence of occurrence of these complexities occurring on both the echoes due to the same travelling disturbances, it is concluded that the disturbances travel from north to south, a result which is in agreement with a similar observation made earlier at this station by the authors (RAO and RAO, 1953, 1954) using a simple c.w. method.

Occurrence of travelling disturbances during night-time has not yet been reported by earlier investigators. The authors have noticed that the travelling disturbances were observed at this station during the earlier part of the night. In all such cases it is found that the separation of $o$ and $x$ rays is small. Fig. 1(f) shows one such observation taken during the night with a $y$ type complexity. It is interesting to note that the $o$ ray and the complexity disappear abruptly at 1915 hours.

*Acknowledgement*—The authors are indebted to the Council of Scientific and Industrial Research (India) for financial support of these investigations.

E. BHAGIRATHA RAO
B. RAMACHANDRA RAO

*Ionospheric Laboratories*
*Andhra University*
*Waltair, India*

## REFERENCES


| | | |
|---|---|---|
| HON YUNG SEN | 1949 | *J. Geophys. Res.* **54,** 363. |
| MUNRO G. M. | 1950 | *Proc. Roy. Soc.* A **202,** 208. |
| MUNRO G. M. | 1953 | *Proc. Roy. Soc.* A **219,** 447. |
| MUNRO G. M. and HEISLER L. M. | 1956 | *Aust. J. Phys.* **9,** 343, 359. |
| OSBORNE B. W. | 1951 | *J. Atmosph. Terr. Phys.* **2,** 66. |
| RAO B. R. and RAO E. B. | 1953 | *Nature, Lond.* **172,** 819. |
| | 1954 | *J. Sci. Industr. Res.* **13-A,** 462. |
| SKINNER N. J., BROWN R. A. and WRIGHT R. W. | 1954 | *J. Atmosph. Terr. Phys.* **5,** 92. |

# The reflection of radio waves from an iceshelf

(*Received* 17 *January* 1961)

THE vertical incidence ionospheric measurements made by the Royal Society Expedition to Halley Bay, Antarctica, may be used to give interesting information about the radio frequency properties of the iceshelf on which the Base is situated.

The total loss, $L_n + L_g$, incurred when a signal is reflected from the ionosphere and from the ground can be determined with high accuracy from the ratio of the amplitudes of a radio frequency pulse reflected once and many times between the ionosphere and the ground (PIGGOTT, 1953). If this loss is measured on two frequencies reflected from the $F$-layer at times when the ionospheric absorption $L_n$ is varying slowly with time, it is possible to deduce $L_g$, and hence the ground reflection coefficient $g$, with considerable accuracy.

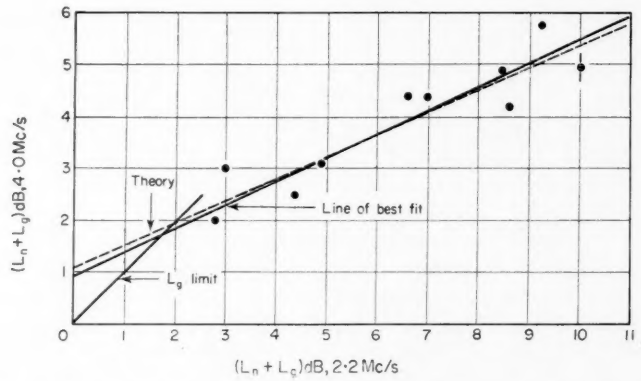


Fig. 1. Determination of the ground loss, $L_g$, from measurements of total absorption made in winter at Halley Bay. The line – – – gives the variation expected if the ionospheric loss is due to a non-deviative absorbing layer.

The method is shown in Fig. 1. The most probable value of the ground reflection loss $L_g$ is given by the interception of the line "$L_g$ limit" and the line of best fit. Clearly the ground reflection coefficient is about 0·8 ($L_g = 2$ dB). A reflection coefficient of this order implies that the dielectric constant of the reflecting surface is near 100. The dielectric constant for a snow or ice sheet lies between 1 and 3. We conclude that the reflection actually occurs at the interface between the sea and the iceshelf and not at the upper interface between the iceshelf and the air. The sea is, of course, an almost perfect reflector at our frequencies.

If the effective reflecting surface is really a considerable distance below the electrical centre of the aerials used at Halley Bay, the power radiated upwards should show variations with frequency whose periodicity is determined by the virtual depth, at radio frequency, of the main reflecting interface. These variations can be detected when the signals reflected from the ionosphere are weak, for example, when the echoes are spread over a wide range of heights. The minima in radiated power show up as small gaps in the traces. These are found at the same radio frequencies on different occasions.

At Halley Bay more than half the predicted minima can be detected in the frequency band 1–11 Mc/s. The mean frequency interval corresponds to a period of 0·6 Mc/s. This shows that the effective depth of the interface which gives the ground reflection is 250 m. When allowance is made for the probable values of the dielectric constant at different depths, calculated from the density of the snow and ice (KUROIWA, 1954), the depth of the

interface is found to be close to 150 m. The depth found by assuming that the iceshelf is in hydrostatic equilibrium, 143 $\pm$ 20 m (MACDOWALL, 1960), agrees well with our value.

We may note that the ground losses given by $L_g$ can be ascribed to the partial reflection coefficient of the air to snow interface. The actual reflection coefficient $\rho_g = 0{\cdot}8$ corresponds to a partial reflection coefficient for this interface of 0·1 (20 dB), a value which is consistent with the snow density near the surface.

*Acknowledgements*—This work forms part of the programme of the British National Committee for the I.G.Y. and the D.S.I.R. Radio Research Station, Slough, and is published by permission of the Director of Radio Research.

W. R. PIGGOTT
L. W. BARCLAY

*D.S.I.R. Radio Research Station*
*Slough, Bucks.**

REFERENCES

KUROIWA D. 1954 U.G.G.I. Assoc. Internat. d'Hydrologie Scientifique. Proc. Gen. Ass. Rome Vol. IV, p. 52.

MACDOWALL J. 1960 *Proc. Roy. Soc.* A **256,** 197.

PIGGOTT W. R. 1953 *Proc. Instn Elect. Engrs* III **100,** 61.

* Official communication.

VOL.
20
961

Journal of Atmospheric and Terrestrial Physics, 1961, Vol. 20, p. 300. Pergamon Press Ltd. Printed in Northern Ireland

# LETTER TO THE EDITOR

## Discussion of the paper "Observed variations in the amplitude scintillations of the Cassiopeia (23N5A) source" by H. J. A. Chivers

(*Received* 27 *December* 1960)

In the course of a discussion on the variation of scintillation rate with position of the radio source Cassiopeia A, Chivers (1960) quotes the earlier work of Reid (1957). The purpose of this communication is to point out an apparent misconception regarding this work.

According to Chivers, Reid concluded that the ionospheric irregularities responsible for scintillation showed a mean drift around the dip-pole because, at two widely separated sites, the maximum scintillation rate occurred when the star was at the same bearing as that of the dip-pole. In fact, even a casual examination of Reid's paper shows that this was not the reason for the conclusion. In that paper an attempt was made to explain the entire shape of the curve of scintillation rate verses sidereal time on the basis of a very simple and admittedly crude model. It was on the basis of this model that the conclusion was drawn regarding the drift of irregularities around the dip-pole. The bearing of the source at maximum scintillation rate was not mentioned.

Chivers further states that there appears to be an error in Reid's work concerning the quoted azimuth of the dip-pole. The author is at a loss to find any "azimuth" for the dip-pole quoted in Reid's paper. The only statement made in this regard is that the average direction of drift at the two sites is roughly perpendicular to the direction of the horizontal component of the local magnetic field, and that normals drawn to the directions of drift at the two sites intersect in a point having approximate geographic coordinates of 70°N, 97°W. This point is much closer to the dip-pole (74°N, 99°W) than to the geomagnetic pole (79°N, 69°W).

G. C. Reid

*Defence Research Telecommunications Establishment*
*Shirley Bay, Ottawa, Canada*

References

Chivers H. J. A. 1960 *J. Atmosph. Terr. Phys.* **19,** 54.
Reid G. C. 1957 *Canad. J. Phys.* **35,** 1004.



VOL. 20 1961

VOL.
20
961